

STIC-Biotech/ChemLib

100452

From: Chan, Christina
Sent: Monday, August 04, 2003 4:40 PM
To: VanderVegt, F. Pierre; STIC-Biotech/ChemLib
Subject: RE: RUSH sequence search needed 09/883,727
Importance: High

Please ~~rush~~ Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: VanderVegt, F. Pierre
Sent: Monday, August 04, 2003 3:16 PM
T: Chan, Christina
Subject: FW: RUSH sequence search needed 09/883,727

Due to a need to move on to additional species in this examination, I need to get a RUSH search done for the following two sequences. An interference search should be included.

Please search SEQ ID NO: 53

and amino acids 64-68 of SEQ ID NO:1 (GRLGC).

Thanks,

F. Pierre (Petr) VanderVegt
Art Unit 1644
9A16
mailbox 9E12

RECEIVED
AUG - 4 2003
(STIC)

Searcher: 11. 34.578
Phone: _____
Location: _____
Date Picked Up: 8/5/03
Date Completed: 8/5/03
Searcher Prep/Review: 4
Clerical: 5
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Access DB# 10045 9

Scientific and Technical Information Center

Requester's Full Name _____ Examiner # _____ Date: _____
 An Unit _____ Phone Number 30 _____ Serial Number _____
 Mail Box and Bldg. Room Location _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the electron species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Jan Detaval
 Reference Librarian
 Chemistry & Chemical Library
 (202) 508-7038
 jan.detaval@nist.gov

STAFF USE ONLY

Searcher: <u>[Signature]</u>	Type of Search	Vendors and cost where applicable
Searcher Phone # <u>4498</u>	NA Sequence (#) _____	STN _____
Searcher Location _____	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Date Searcher Received <u>6/16/03</u>	Structure (#) _____	Onsite Order _____
Date Time Stamp <u>5/15/03</u>	Bibliographic _____	_____
Searcher Prep & Review Time _____	Litigation _____	Books & News _____
Prep Time <u>10</u>	Fulltext _____	Sequences & Systems <input checked="" type="checkbox"/>
Review Time <u>15</u>	Patent Family _____	USNA Inquiries _____
	Other _____	_____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:36:20 ; Search time 10.4286 Seconds
(without alignments)
76.102 Million cell updates/sec

Title: US-09-883-727A-1_COPY_64_68

Perfect score: 33
Sequence: 1 CRLGC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	6	AAE20046	Complement C1s cat
2	33	100.0	7	AAE20047	Complement C1s cat
3	33	100.0	8	AAE20048	Complement C1s cat
4	33	100.0	9	AAE20049	Complement C1s cat
5	33	100.0	10	AAE20050	Complement C1s cat
6	33	100.0	11	AAE20051	Complement C1s cat
7	33	100.0	12	AAE20052	Complement C1s cat
8	33	100.0	13	AAE20053	Complement C1s cat
9	33	100.0	14	AAE20054	Complement C1s cat

```
10 33 100.0 15 23 AAE20055 Complement C1s cat
11 33 100.0 16 23 AAE20056 Complement C1s cat
12 33 100.0 16 23 ABB88255 C marmoreus mu-con
13 33 100.0 16 23 ABB88256 C marmoreus mu-con
14 33 100.0 16 23 ABB88257 C marmoreus mu-con
15 33 100.0 16 23 ABB88458 C marmoreus mu-con
16 33 100.0 16 23 ABB88459 C marmoreus mu-con
17 33 100.0 16 23 ABB88460 C marmoreus mu-con
18 33 100.0 17 23 AAE20057 Complement C1s cat
19 33 100.0 18 23 AAE20058 Complement C1s cat
20 33 100.0 19 23 AAE20059 Complement C1s cat
21 33 100.0 20 23 AAE20060 Complement C1s cat
22 33 100.0 21 23 AAE20061 Complement C1s cat
23 33 100.0 22 23 AAE20062 Complement C1s cat
24 33 100.0 23 23 AAE20063 Complement C1s cat
25 33 100.0 29 23 AAE20068 Complement C1s cat
26 33 100.0 30 23 AAE20000 Complement C1s cat
27 33 100.0 30 23 AAE20006 Complement C1s cat
28 33 100.0 30 23 AAE20007 Complement C1s cat
29 33 100.0 30 23 AAE20013 Complement C1s cat
30 33 100.0 30 23 AAE20018 Complement C1s cat
31 33 100.0 30 23 AAE20020 Complement C1s cat
32 33 100.0 30 23 AAE20021 Complement C1s cat
33 33 100.0 30 23 AAE20023 Complement C1s cat
34 33 100.0 30 23 AAE20024 Complement C1s cat
35 33 100.0 30 23 AAE20027 Complement C1s cat
36 33 100.0 30 23 AAE20029 Complement C1s cat
37 33 100.0 30 23 AAE20032 Complement C1s cat
38 33 100.0 30 23 AAE20034 Complement C1s cat
39 33 100.0 30 23 AAE20035 Complement C1s cat
40 33 100.0 30 23 AAE20037 Complement C1s cat
41 33 100.0 30 23 AAE20038 Complement C1s cat
42 33 100.0 30 23 AAE20041 Complement C1s cat
43 33 100.0 30 23 AAE20045 Complement C1s cat
44 33 100.0 53 22 AAU59535 Propionibacterium
45 33 100.0 59 22 AAU66795 Propionibacterium
```

ALIGNMENTS

RESULT 1
AAE20046
ID AAE20046 standard; peptide; 6 AA.

XX AC AAE20046;

XX DT 18-JUN-2002 (first entry)

XX XX Complement C1s catalytic site-directed moiety, peptide #47.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
XX KW Preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
XX KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
XX KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212948P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX XX WPI; 2002-241177/29.

XX XX

PT New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis
 XX
 PS Disclosure; Page 20; 99pp; English.
 XX
 CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 33; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 1 CRLGC 5
 RESULT 2
 AAE20047
 ID AAE20047 standard; peptide; 7 AA.
 AC AAE20047;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Complement C1s catalytic site-directed moiety, peptide #48.
 KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.
 XX
 PN WO200198365-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 18-JUN-2001; 2001WO-US19405.
 XX
 PP 21-JUN-2000; 2000US-212998P.
 XX
 PR (ZYMO) ZYMOGENETICS INC.
 XX
 PA West RR, Sheppard PO, Fox BA;
 XX
 PI WPI; 2002-241177/29.
 XX
 PS New complement C1s inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 XX sepsis
 XX
 PS Disclosure; Page 20; 99pp; English.
 XX
 CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 33; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 1 CRLGC 5
 RESULT 3
 AAE20048
 ID AAE20048 standard; peptide; 8 AA.
 AC AAE20048;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Complement C1s catalytic site-directed moiety, peptide #49.
 KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.
 XX
 PN WO200198365-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 18-JUN-2001; 2001WO-US19405.
 XX
 PP 21-JUN-2000; 2000US-212998P.
 XX
 PR (ZYMO) ZYMOGENETICS INC.
 XX
 PA West RR, Sheppard PO, Fox BA;
 XX
 PI WPI; 2002-241177/29.
 XX
 PS New complement C1s inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 XX sepsis
 XX
 PS Disclosure; Page 20; 99pp; English.
 XX
 CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 33; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 Db 1 CRLGC 5

RESULT 4

AAE20049
 ID AAE20049 standard; peptide; 9 AA.

XX AC AAE20049;

XX DT 18-JUN-2002 (first entry)

XX DE Complement C1s catalytic site-directed moiety, peptide #50.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX PT New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis

XX PS Disclosure; Page 20; 99pp; English.

XX CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejectin, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 33; DB 23; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 1 CRLGC 5

RESULT 5

AAE20050

ID AAE20050 standard; peptide; 10 AA.

XX AC AAE20050;

XX DT 18-JUN-2002 (first entry)

XX

DE Complement C1s catalytic site-directed moiety, peptide #51.

XX

KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX PT New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis

XX PS Disclosure; Page 20; 99pp; English.

XX

CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejectin, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 33; DB 23; Length 10;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 1 CRLGC 5

RESULT 6

AAE20051

ID AAE20051 standard; peptide; 11 AA.

XX AC AAE20051;

XX DT 18-JUN-2002 (first entry)

XX DE Complement C1s catalytic site-directed moiety, peptide #52.

XX

KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX Sequence 13 AA;

Query Match 100.0%; Score 33; DB 23; Length 13;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 1 CRLGC 5

RESULT 9

AAE20054
 ID AAE20054 standard; peptide; 14 AA.

XX AAE20054;

DT 18-JUN-2002 (first entry)

DE Complement C1s catalytic site-directed moiety, peptide #55.

CC Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

OS Unidentified.

XX WO200198365-A2.

PN 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

CC New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
 PT sepsis -

PS Disclosure; Page 20; 99pp; English.

CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX Sequence 14 AA;

Query Match 100.0%; Score 33; DB 23; Length 14;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 1 CRLGC 5

RESULT 10

AAE20055
 ID AAE20055 standard; peptide; 15 AA.

XX AAE20055;

DT 18-JUN-2002 (first entry)

DE Complement C1s catalytic site-directed moiety, peptide #56.

CC Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

OS Unidentified.

XX WO200198365-A2.

PN 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

CC New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
 PT sepsis -

PS Disclosure; Page 20; 99pp; English.

CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX Sequence 15 AA;

Query Match 100.0%; Score 33; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 1 CRLGC 5

RESULT 11

AAE20056
 ID AAE20056 standard; peptide; 16 AA.

XX AAE20056;

DT 18-JUN-2002 (first entry)
 XX Complement CIs catalytic site-directed moiety, peptide #57.
 DE
 XX
 KW Complement CIs; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.
 XX
 XX WO200198365-A2.
 XX
 XX 27-DEC-2001.
 PD
 XX
 XX 18-JUN-2001; 2001WO-US19405.
 PF
 XX
 XX 21-JUN-2000; 2000US-212998P.
 PR
 XX
 XX (2YMO) ZYMOGENETICS INC.
 PA
 XX
 XX West RR, Sheppard PO, Fox BA;
 PI
 XX WPI; 2002-241177/29.
 DR
 XX
 XX New complement CIs inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis
 XX
 XX Disclosure; Page 20; 99pp; English.
 PS
 XX
 XX The invention relates to a polypeptide that inhibits complement CIs. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate CIs.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC peptide sequence is complement CIs catalytic site-directed moiety.
 XX
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 33; DB 23; Length 16;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 1 CRLGC 5
 RESULT 12
 ABB88255
 ID ABB88255 standard; Peptide; 16 AA.
 XX
 XX ABB88255;
 AC
 XX
 XX 24-MAY-2002 (first entry)
 DT
 XX
 XX C marmoreus mu-conopeptide U019.
 DE
 XX
 XX Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;
 KW anaesthetic; neuromuscular blocking agent; neuroprotective; pain;
 KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;
 KW cardiovascular; vasotropic; cardiant; tranquilizer; antimigraine;
 KW neurodegenerative disease; neuromuscular disorder.
 XX
 OS Conus marmoreus.
 XX

PH Key Location/Qualifiers
 FT Misc-difference 4 /label= Pro, OTHER
 FT /note= "OTHER=hydroxy-Pro"
 FT Misc-difference 13 /label= Pro, OTHER
 FT /note= "OTHER=hydroxy-Pro"
 XX
 PN WO200207678-A2.
 XX
 PD 31-JAN-2002.
 XX
 XX 23-JUL-2001; 2001WO-US23125.
 PF
 XX
 XX 21-JUL-2000; 2000US-219619P.
 PR
 XX 03-NOV-2000; 2000US-245157P.
 PR
 XX 29-JAN-2001; 2001US-264319P.
 PR
 XX 21-MAR-2001; 2001US-277270P.
 PR
 XX (UTAH) UNIV UTAH RES FOUND.
 PA (COGN-) COGNETIX INC.
 PA
 XX Olivera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;
 PI Jacobsen R, Jones RM, Cartier GE, Shen GS;
 PI WPI; 2002-217020/27.
 XX
 XX New mu-conopeptides useful for treating disorders associated with
 PT voltage-gated sodium channels, e.g. stroke or pain, as neuromuscular
 PT blocking agents, as local anesthetic agents, as analgesic agents and as
 PT neuroprotective agents
 XX
 XX Claim 1; Page 50; 231pp; English.
 PS
 XX The present invention relates to mu-conopeptides derived from snails,
 CC which can be in the treatment of disorders associated with voltage-gated
 CC ion channels. These may include neurodegenerative disorders such as
 CC ankyrotrophic lateral sclerosis, neuromuscular disorders, post-operative or
 CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,
 CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,
 CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia
 CC or ischaemia which typically follows stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for
 CC providing muscle relaxation, treating essential blepharospasm and other
 CC forms of focal dystonia, and for anti-wrinkle use. The present sequence
 CC is a mu-conopeptide of the invention.
 XX
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 33; DB 23; Length 16;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 7 CRLGC 11
 RESULT 13
 ABB88256
 ID ABB88256 standard; Peptide; 16 AA.
 XX
 XX ABB88256;
 AC
 XX
 XX 24-MAY-2002 (first entry)
 DT
 XX
 XX C marmoreus mu-conopeptide U20.
 DE
 XX
 XX Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;
 KW anaesthetic; neuromuscular blocking agent; neuroprotective; pain;
 KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;
 KW cardiovascular; vasotropic; cardiant; tranquilizer; antimigraine;
 KW Conus marmoreus.

KW neurodegenerative disease; neuromuscular disorder.

XX Conus marmoreus.

XX Key Location/Qualifiers

FT Misc-difference 4 /label= Pro, OTHER

FT /note= "OTHER=hydroxy-Pro"

FT Misc-difference 13

FT /label= Pro, OTHER

FT /note= "OTHER=hydroxy-Pro"

XX WO200207678-A2.

XX 31-JAN-2002.

XX 23-JUL-2001; 2001WO-US23125.

XX 21-JUL-2000; 2000US-219619P.

XX 03-NOV-2000; 2000US-245157P.

XX 29-JAN-2001; 2001US-264319P.

XX 21-MAR-2001; 2001US-277270P.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

PI Olivera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;

PI Jacobsen R, Jones RM, Cartier GE, Shen GS;

XX WPI; 2002-217020/27.

XX New mu-conopeptides useful for treating disorders associated with

PT voltage-gated sodium channels, e.g. stroke or pain, as neuromuscular

PT blocking agents, as local anesthetic agents, as analgesic agents and as

PT neuroprotective agents

XX Claim 1; Page 50; 231pp; English.

XX The present invention relates to mu-conopeptides derived from snails,

CC which can be in the treatment of disorders associated with voltage-gated

CC ion channels. These may include neurodegenerative disorders such as

CC amyotrophic lateral sclerosis, neuromuscular disorders, post-operative or

CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,

CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,

CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia

CC or ischaemia which typically follows stroke, cerebrovascular accident,

CC brain or spinal cord trauma, myocardial infarct, physical trauma,

CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for

CC providing muscle relaxation, treating essential blepharospasm and other

CC forms of focal dystonia, and for anti-wrinkle use. The present sequence

CC is a mu-conopeptide of the invention.

XX Sequence 16 AA;

SQ Query Match 100.0%; Score 33; DB 23; Length 16;

XX Best Local Similarity 100.0%; Pred. No. 52;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

DB 7 CRLGC 11

RESULT 14

ABB88257

ID ABB88257 standard; Peptide; 16 AA.

XX ABB88257;

AC ABB88257;

DT 24-MAY-2002 (first entry)

XX C marmoreus mu-conopeptide U022.

DE C marmoreus mu-conopeptide U022.

XX

KW Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;

KW anaesthetic; neuromuscular blocking agent; neuroprotective; pain;

KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;

KW cardiovascular; vasotropic; cardiac; tranquilizer; antimigraine;

XX neurodegenerative disease; neuromuscular disorder.

OS Conus marmoreus.

XX Key Location/Qualifiers

FT Misc-difference 4 /label= Pro, OTHER

FT /note= "OTHER=hydroxy-Pro"

FT Misc-difference 13

FT /label= Pro, OTHER

FT /note= "OTHER=hydroxy-Pro"

XX WO200207678-A2.

XX 31-JAN-2002.

XX 23-JUL-2001; 2001WO-US23125.

XX 21-JUL-2000; 2000US-219619P.

XX 03-NOV-2000; 2000US-245157P.

XX 29-JAN-2001; 2001US-264319P.

XX 21-MAR-2001; 2001US-277270P.

XX (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

PI Olivera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;

PI Jacobsen R, Jones RM, Cartier GE, Shen GS;

XX WPI; 2002-217020/27.

XX New mu-conopeptides useful for treating disorders associated with

PT voltage-gated sodium channels, e.g. stroke or pain, as neuromuscular

PT blocking agents, as local anesthetic agents, as analgesic agents and as

PT neuroprotective agents

XX Claim 1; Page 50; 231pp; English.

XX The present invention relates to mu-conopeptides derived from snails,

CC which can be in the treatment of disorders associated with voltage-gated

CC ion channels. These may include neurodegenerative disorders such as

CC amyotrophic lateral sclerosis, neuromuscular disorders, post-operative or

CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,

CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,

CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia

CC or ischaemia which typically follows stroke, cerebrovascular accident,

CC brain or spinal cord trauma, myocardial infarct, physical trauma,

CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for

CC providing muscle relaxation, treating essential blepharospasm and other

CC forms of focal dystonia, and for anti-wrinkle use. The present sequence

CC is a mu-conopeptide of the invention.

XX Sequence 16 AA;

SQ Query Match 100.0%; Score 33; DB 23; Length 16;

XX Best Local Similarity 100.0%; Pred. No. 52;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

DB 7 CRLGC 11

RESULT 15

ABB88458

ID ABB88458 standard; Peptide; 16 AA.

XX ABB88458;

AC ABB88458;

XX

DT 24-MAY-2002 (first entry)
 XX
 DE C marmoreus mu-conopeptide SEQ ID NO: 441.
 XX
 KW Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;
 KW anaesthetic; neuromuscular blocking agent; neuroprotective; pain;
 KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;
 KW cardiovascular; vasotropic; cardiant; tranquilizer; antimigraine;
 KW neurodegenerative disease; neuromuscular disorder.
 XX
 OS Conus marmoreus.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 13
 FT /label= Xaa
 FT /note= "Xaa-hydroxy-Pro"
 XX
 PN WO200207678-A2.
 XX
 PD 31-JAN-2002.
 XX
 PD 23-JUL-2001; 2001WO-US23125.
 XX
 PF 21-JUL-2000; 2000US-219619P.
 PR 03-NOV-2000; 2000US-245157P.
 PR 29-JAN-2001; 2001US-264319P.
 PR 21-MAR-2001; 2001US-277270P.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 PA (COGN-) COGNETIX INC.
 XX
 PI Olivera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;
 PI Jacobsen R, Jones RM, Cartier GE, Shen GS;
 XX
 DR WPI; 2002-217020/27.
 XX
 PT New mu-conopeptides useful for treating disorders associated with
 PT voltage-gated sodium channels, e.g. stroke or pain, as neuromuscular
 PT blocking agents, as local anesthetic agents, as analgesic agents and as
 PT neuroprotective agents .
 XX
 FS Claim 1; Page 83; 23lpp; English.
 XX
 CC The present invention relates to mu-conopeptides derived from snails,
 CC which can be in the treatment of disorders associated with voltage-gated
 CC ion channels. These may include neurodegenerative disorders such as
 CC anyotrophic lateral sclerosis, neuromuscular disorders, post-operative or
 CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,
 CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,
 CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia
 CC or ischaemia which typically follows stroke, cerebrovascular accident,
 CC brain or spinal cord trauma, myocardial infarct, physical trauma,
 CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for
 CC providing muscle relaxation, treating essential blepharospasm and other
 CC forms of focal dystonia, and for anti-wrinkle use. The present sequence
 CC is a mu-conopeptide of the invention.
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 33; DB 23; Length 16;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 Db 7 CRLGC 11

Search completed: August 5, 2003, 09:39:26
 Job time : 10.4286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:38:10 ; Search time 4.14286 Seconds
(without alignments)
51.065 Million cell updates/sec

Title: US-09-883-727a-1_COPY_64_68

Perfect score: 33

Sequence: 1 CRLGC 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/prodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2.6/prodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2.6/prodata/1/1aa/6A-COMB.pep:*
- 4: /cgn2.6/prodata/1/1aa/6B-COMB.pep:*
- 5: /cgn2.6/prodata/1/1aa/FCUS-COMB.pep:*
- 6: /cgn2.6/prodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	246	4	US-09-252-991A-23684
2	33	100.0	1665	3	US-09-858-664A-2
3	31	93.9	13	3	US-09-258-754-115
4	31	93.9	13	3	US-09-042-107-115
5	31	93.9	373	4	US-09-134-001C-4798
6	31	93.9	2254	2	US-08-286-819A-28
7	31	93.9	2254	3	US-08-980-357-28
8	30	90.9	1094	2	US-08-680-326-40
9	29	87.9	1181	4	US-09-252-991A-18714
10	29	87.9	1286	4	US-09-170-496D-291
11	28	84.8	13	4	US-09-089-878-2
12	28	84.8	39	2	US-08-867-087B-50
13	28	84.8	70	4	US-08-936-165A-458
14	28	84.8	103	4	US-09-252-991A-27147
15	28	84.8	133	4	US-09-252-991A-30594
16	28	84.8	134	4	US-09-252-991A-25718
17	28	84.8	148	4	US-09-252-991A-26878
18	28	84.8	154	4	US-09-252-991A-16837
19	28	84.8	177	4	US-09-149-476-565
20	28	84.8	179	4	US-09-252-991A-27201
21	28	84.8	184	4	US-09-252-991A-31672
22	28	84.8	189	2	US-08-861-269-7
23	28	84.8	189	2	US-09-134-596-7
24	28	84.8	189	3	US-09-293-273-7
25	28	84.8	209	4	US-09-252-991A-17227
26	28	84.8	217	4	US-09-252-991A-19838
27	28	84.8	217	4	US-09-252-991A-25975

28	84.8	253	2	US-08-592-214A-4	Sequence 4, Appl1
29	84.8	253	3	US-08-659-188-4	Sequence 4, Appl1
30	28	253	3	US-08-655-227-4	Sequence 4, Appl1
31	28	253	3	US-08-655-241-4	Sequence 4, Appl1
32	28	253	3	US-09-149-976-4	Sequence 4, Appl1
33	28	253	3	US-09-398-326-4	Sequence 4, Appl1
34	28	255	2	US-08-576-156-2	Sequence 2, Appl1
35	28	256	3	US-08-659-188-2	Sequence 2, Appl1
36	28	256	3	US-08-552-227-2	Sequence 2, Appl1
37	28	256	3	US-08-655-241-2	Sequence 2, Appl1
38	28	256	4	US-09-398-326-2	Sequence 2, Appl1
39	28	264	4	US-08-469-260A-75	Sequence 75, Appl1
40	28	264	4	US-09-252-991A-24262	Sequence 24262, A
41	28	264	4	US-08-488-446-75	Sequence 75, Appl1
42	28	264	4	US-08-467-344A-75	Sequence 75, Appl1
43	28	273	4	US-09-149-476-476	Sequence 476, App
44	28	300	2	US-08-592-214A-2	Sequence 2, Appl1
45	28	300	3	US-09-149-976-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-252-991A-23684
; Sequence 23684, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33143
; SEQ ID NO 23684
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23684

Query Match 100.0%; Score 33; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CRLGC 5
Db	124	CRLGC 128

RESULT 2
US-09-858-664A-2
; Sequence 2, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-858-664A-2

Query Match 100.0%; Score 33; DB 4; Length 1665;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
Db 4 CRLGC 8

RESULT 3

US-09-258-754-115
; Sequence 115, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Diptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 115
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-115

Query Match 93.9%; Score 31; DB 3; Length 13;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
Db 9 CRMG 13

RESULT 4

US-09-042-107-115
; Sequence 115, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
; TITLE OF INVENTION: Tissues
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 115
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-115

Query Match 93.9%; Score 31; DB 3; Length 13;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 9 CRMG 13
 |||||

RESULT 5

US-09-134-001C-4798
; Sequence 4798, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4798
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4798

Query Match 93.9%; Score 31; DB 4; Length 373;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
Db 136 CRIG 140

RESULT 6

US-08-286-819A-28
; Sequence 28, Application US/08286819A
; Patent No. 5871910
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DUKTA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPETIDES, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,819A
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/174,682
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,146
; FILING DATE: 10-AUG-1992

CLASSIFICATION: 435
PRIOR APPLICATION DATA: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-28

Query Match . 93.9%; Score 31; DB 2; Length 2254;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRIGC 5
Db 1367 CRIGC 1371

RESULT 7
US-08-980-357-28
Sequence 28, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-28

Query Match . 93.9%; Score 31; DB 3; Length 2254;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRIGC 5
Db 1367 CRIGC 1371

RESULT 8
US-08-680-326-40
Sequence 40, Application US/08680326
Patent No. 5925733
GENERAL INFORMATION:
APPLICANT: ROSE, TIMOTHY M.
APPLICANT: BOSCH, MARNIX
APPLICANT: STRAND, KURT
APPLICANT: TODARO, GEORGE J.
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
TITLE OF INVENTION: FIBROMATOSIS
NUMBER OF SEQUENCES: 152
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,326
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1094 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-680-326-40

Query Match 90.9%; Score 30; DB 2; Length 1094;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 92 CRVGC 96

RESULT 9

US-09-252-991A-18714
; Sequence 18714, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18714
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18714

Query Match 87.9%; Score 29; DB 4; Length 1181;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 126 CRVGC 130

RESULT 10

US-09-170-496D-291
; Sequence 291, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 291
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-291

Query Match 87.9%; Score 29; DB 4; Length 1286;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 306 COLGC 310

RESULT 11

US-09-089-878-2
; Sequence 2, Application US/09089878
; Patent No. 6458528
; GENERAL INFORMATION:
; APPLICANT: Groat, Randall G.
; APPLICANT: O'Connor, Thomas P.
; TITLE OF INVENTION: DIAGNOSIS OF FELINE IMMUNODEFICIENCY VIRUS INFECTION
; FILE REFERENCE: 00088/111001
; CURRENT APPLICATION NUMBER: US/09/089,878
; CURRENT FILING DATE: 1998-06-03
; EARLIER APPLICATION NUMBER: US 60/085,615
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Feline immunodeficiency virus
US-09-089-878-2

Query Match 84.8%; Score 28; DB 4; Length 13;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CELGC 5

RESULT 12

US-08-867-087B-50
; Sequence 50, Application US/08867087B
; Patent No. 5990386
; GENERAL INFORMATION:
; APPLICANT: An, Gynheung
; TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
; FILE REFERENCE: 70
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
; ADDRESSEE: Winston, LLP
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,087B
; FILING DATE: June 2, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/323,449
; FILING DATE: October 14, 1994
; APPLICATION NUMBER: U.S. 08/485,981
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Alan E.
; REGISTRATION NUMBER: 35,123
; REFERENCE/DOCKET NUMBER: 4630-47071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
US-08-867-087B-50

Query Match 84.8%; Score 28; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 32 CNLGC 36

RESULT 13
US-08-936-165A-458
Sequence 458, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 458:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-458

Query Match 84.8%; Score 28; DB 4; Length 70;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRLGC 5

DB 58 CELGC 62

RESULT 14

US-09-252-991A-27147
Sequence 27147, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.135
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27147
LENGTH: 103
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27147

Query Match 84.8%; Score 28; DB 4; Length 103;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 59 CRTGC 63

RESULT 15

US-09-252-991A-30594
Sequence 30594, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30594
LENGTH: 133
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30594

Query Match 84.8%; Score 28; DB 4; Length 133;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 58 CRAGC 62

Search completed: August 5, 2003, 09:42:02
Job time : 6.14286 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:41:00 ; Search time 9.28571 Seconds
(without alignments)
63.948 Million cell updates/sec

Title: US-09-883-727A-1_COPY_64_68
Perfect score: 33
Sequence: 1 CRLGC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	6	10 US-09-883-727A-99	Sequence 99, Appl
2	33	100.0	7	10 US-09-883-727A-100	Sequence 100, App
3	33	100.0	8	10 US-09-883-727A-101	Sequence 101, App
4	33	100.0	9	10 US-09-883-727A-102	Sequence 102, App
5	33	100.0	10	10 US-09-883-727A-103	Sequence 103, App
6	33	100.0	11	10 US-09-883-727A-104	Sequence 104, App
7	33	100.0	12	10 US-09-883-727A-105	Sequence 105, App
8	33	100.0	13	10 US-09-883-727A-106	Sequence 106, App
9	33	100.0	14	10 US-09-883-727A-107	Sequence 107, App
10	33	100.0	15	10 US-09-883-727A-108	Sequence 108, App
11	33	100.0	16	10 US-09-883-727A-109	Sequence 109, App
12	33	100.0	16	11 US-09-910-009A-170	Sequence 170, App
13	33	100.0	16	11 US-09-910-009A-171	Sequence 171, App
14	33	100.0	16	11 US-09-910-009A-172	Sequence 172, App
15	33	100.0	16	11 US-09-910-009A-441	Sequence 441, App

16	33	100.0	16	11	US-09-910-009A-442	Sequence 442, App
17	33	100.0	16	11	US-09-910-009A-443	Sequence 443, App
18	33	100.0	17	10	US-09-883-727A-110	Sequence 110, App
19	33	100.0	18	10	US-09-883-727A-111	Sequence 111, App
20	33	100.0	19	10	US-09-883-727A-112	Sequence 112, App
21	33	100.0	20	10	US-09-883-727A-113	Sequence 113, App
22	33	100.0	21	10	US-09-883-727A-114	Sequence 114, App
23	33	100.0	22	10	US-09-883-727A-115	Sequence 115, App
24	33	100.0	23	10	US-09-883-727A-116	Sequence 116, App
25	33	100.0	29	10	US-09-883-727A-121	Sequence 121, App
26	33	100.0	30	10	US-09-883-727A-53	Sequence 53, Appl
27	33	100.0	30	10	US-09-883-727A-59	Sequence 59, Appl
28	33	100.0	30	10	US-09-883-727A-60	Sequence 60, Appl
29	33	100.0	30	10	US-09-883-727A-66	Sequence 66, Appl
30	33	100.0	30	10	US-09-883-727A-71	Sequence 71, Appl
31	33	100.0	30	10	US-09-883-727A-73	Sequence 73, Appl
32	33	100.0	30	10	US-09-883-727A-74	Sequence 74, Appl
33	33	100.0	30	10	US-09-883-727A-76	Sequence 76, Appl
34	33	100.0	30	10	US-09-883-727A-77	Sequence 77, Appl
35	33	100.0	30	10	US-09-883-727A-80	Sequence 80, Appl
36	33	100.0	30	10	US-09-883-727A-82	Sequence 82, Appl
37	33	100.0	30	10	US-09-883-727A-85	Sequence 85, Appl
38	33	100.0	30	10	US-09-883-727A-87	Sequence 87, Appl
39	33	100.0	30	10	US-09-883-727A-88	Sequence 88, Appl
40	33	100.0	30	10	US-09-883-727A-90	Sequence 90, Appl
41	33	100.0	30	10	US-09-883-727A-91	Sequence 91, Appl
42	33	100.0	30	10	US-09-883-727A-94	Sequence 94, Appl
43	33	100.0	30	10	US-09-883-727A-98	Sequence 98, Appl
44	33	100.0	122	10	US-09-883-727A-1	Sequence 1, Appl
45	33	100.0	211	11	US-09-764-891-4967	Sequence 4967, Ap

ALIGNMENTS

RESULT 1
US-09-883-727A-99
; Sequence 99, Application US/C9883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul C.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-99

Query Match 100.0%; Score 33; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 1 CRLGC 5

RESULT 2
US-09-883-727A-100
; Sequence 100, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 100
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-100

Query Match 100.0%; Score 33; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CRLGC 5

RESULT 3

US-09-883-727A-101
; Sequence 101, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-101

Query Match 100.0%; Score 33; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CRLGC 5

RESULT 4

US-09-883-727A-102
; Sequence 102, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 102
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-102

Query Match 100.0%; Score 33; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CRLGC 5

RESULT 5

US-09-883-727A-103
; Sequence 103, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 103
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-103

Query Match 100.0%; Score 33; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CRLGC 5

RESULT 6

US-09-883-727A-104
; Sequence 104, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-104

Query Match 100.0%; Score 33; DB 10; Length 11;

Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

RESULT 7

US-09-883-727A-105
; Sequence 105, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 105

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-105

Query Match 100.0%; Score 33; DB 10; Length 12;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

RESULT 8

US-09-883-727A-106

; Sequence 106, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 106

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-106

Query Match 100.0%; Score 33; DB 10; Length 13;

Best Local Similarity 100.0%; Pred. No. 8.7;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

RESULT 9

US-09-883-727A-109

; Sequence 109, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 109

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-109

Query Match 100.0%; Score 33; DB 10; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

US-09-883-727A-107

; Sequence 107, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 107

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-107

Query Match 100.0%; Score 33; DB 10; Length 14;

Best Local Similarity 100.0%; Pred. No. 9.2;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

RESULT 10

US-09-883-727A-108

; Sequence 108, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 108

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-108

Query Match 100.0%; Score 33; DB 10; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

RESULT 11

US-09-883-727A-109

; Sequence 109, Application US/09883727A
; Patent No. US20020102256A1

; GENERAL INFORMATION:

; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Fox, Brian

; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of

; FILE REFERENCE: 00-33

; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 109

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: C1s catalytic site-directed moiety

US-09-883-727A-109

Query Match 100.0%; Score 33; DB 10; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 1 CRLGC 5

FILE REFERENCE: 00-33
CURRENT APPLICATION NUMBER: US/09/883,727A
CURRENT FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 140
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 109
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cls catalytic site-directed moiety
US-09-883-727A-109

Query Match 100.0%; Score 33; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 1 CRLGC 5

RESULT 12

US-09-910-009A-170
Sequence 170, Application US/09910009A
Publication No. US20030050234A1
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Olivera, Baldomero M.
APPLICANT: McIntosh, J. Michael
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Cruz, Lourdes J.
APPLICANT: Shon, Ki-Joon
APPLICANT: Jacobsen, Richard
APPLICANT: Jones, Robert M.
APPLICANT: Cartier, G. Edward
APPLICANT: Shen, Greg S.
APPLICANT: Wagstaff, John D.
TITLE OF INVENTION: Mu-Conopeptides
FILE REFERENCE: 2314-242
CURRENT APPLICATION NUMBER: US/09/910,009A
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 60/219,619
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: US 60/245,157
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/264,319
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: US 60/277,270
NUMBER OF SEQ ID NOS: 520
SOFTWARE: PatentIn version 3.0
SEQ ID NO 170
LENGTH: 16
TYPE: PRT
ORGANISM: Conus marmoreus
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(16)
OTHER INFORMATION: Xaa at residue 4 and 13 is Pro or Hyp

Query Match 100.0%; Score 33; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 7 CRLGC 11

RESULT 13
US-09-910-009A-171
Sequence 171, Application US/09910009A
Publication No. US20030050234A1
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Olivera, Baldomero M.
APPLICANT: McIntosh, J. Michael
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Cruz, Lourdes J.
APPLICANT: Shon, Ki-Joon
APPLICANT: Jacobsen, Richard
APPLICANT: Jones, Robert M.
APPLICANT: Cartier, G. Edward
APPLICANT: Shen, Greg S.
APPLICANT: Wagstaff, John D.
TITLE OF INVENTION: Mu-Conopeptides
FILE REFERENCE: 2314-242
CURRENT APPLICATION NUMBER: US/09/910,009A
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 60/219,619
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: US 60/245,157
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/264,319
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: US 60/277,270
NUMBER OF SEQ ID NOS: 520
SOFTWARE: PatentIn version 3.0
SEQ ID NO 171
LENGTH: 16
TYPE: PRT
ORGANISM: Conus marmoreus
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(16)
OTHER INFORMATION: Xaa at residue 4 and 13 is Pro or Hyp
US-09-910-009A-171

Query Match 100.0%; Score 33; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 7 CRLGC 11

RESULT 14

US-09-910-009A-172
Sequence 172, Application US/09910009A
Publication No. US20030050234A1
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Olivera, Baldomero M.
APPLICANT: McIntosh, J. Michael
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Cruz, Lourdes J.
APPLICANT: Shon, Ki-Joon
APPLICANT: Jacobsen, Richard
APPLICANT: Jones, Robert M.
APPLICANT: Cartier, G. Edward
APPLICANT: Shen, Greg S.
APPLICANT: Wagstaff, John D.
TITLE OF INVENTION: Mu-Conopeptides
FILE REFERENCE: 2314-242
CURRENT APPLICATION NUMBER: US/09/910,009A
CURRENT FILING DATE: 2001-07-23

; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/245,157
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 60/277,270
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 520
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 172
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Xaa at residue 4 and 13 is Pro or Hyp
US-09-910-009A-172

Query Match 100.0%; Score 33; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 7 CRLGC 11

RESULT 15

US-09-910-009A-441
; Sequence 441, Application US/09910009A
; Publication No. US2003005234A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Shen, Greg S.
; APPLICANT: Wagstaff, John D.
; TITLE OF INVENTION: Mu-Conopeptides
; FILE REFERENCE: 2314-242
; CURRENT APPLICATION NUMBER: US/09/910,009A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/245,157
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 60/277,270
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 520
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 441
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Xaa is Hyp
US-09-910-009A-441

Query Match 100.0%; Score 33; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 10;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRLGC 5
|||||
Db 7 CRLGC 11
Search completed: August 5, 2003, 09:52:46
Job time : 9.78571 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:37:45 ; Search time 3.71429 Seconds
(without alignments)
129.458 Million cell updates/sec

Title: US-09-883-727A-1_COPY_64_68

Perfect score: 33

Sequence: 1 CRLGC 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	83	2 C82247	probable proteinase
2	33	100.0	324	2 S75735	probable malate de
3	33	100.0	358	2 D84901	hypothetical prote
4	33	100.0	362	2 B83963	hypothetical prote
5	33	100.0	561	2 AE3094	oxidoreductase Atu
6	33	100.0	561	2 E98192	probable oxidoredu
7	33	100.0	783	2 AH2226	hypothetical prote
8	33	100.0	791	2 AF2235	hypothetical prote
9	33	100.0	1135	1 JQ1928	G2-G1 polypeptid
10	31	93.9	114	2 G90296	hypothetical prote
11	31	93.9	194	2 A82412	hypothetical prote
12	31	93.9	207	2 S71452	trypsin/chymotryps
13	31	93.9	345	2 F97113	probable Fe-S-clus
14	31	93.9	363	2 F69878	conserved hypothet
15	31	93.9	364	2 E89894	conserved hypothet
16	31	93.9	1357	2 S57052	hypothetical prote
17	31	93.9	1518	2 T28880	hypothetical prote
18	30	90.9	116	2 T41597	hypothetical prote
19	30	90.9	118	2 C96746	thionin, 63255-627
20	30	90.9	125	2 S52548	thionin variant Th
21	30	90.9	125	2 S52550	thionin variant Th
22	30	90.9	133	2 S52554	thionin variant Th
23	30	90.9	133	2 S52547	thionin variant Th
24	30	90.9	134	2 S52553	thionin variant Th
25	30	90.9	134	2 S52546	thionin variant Th
26	30	90.9	134	2 F96685	probable thionin F
27	30	90.9	135	2 S52552	thionin variant Th
28	30	90.9	135	2 S52555	thionin variant Th
29	30	90.9	135	2 H84523	probable thionin (

30 30 90.9 303 1 P3BYAM
31 30 90.9 425 1 E70351
32 30 90.9 585 2 H84966
33 30 90.9 637 2 H96592
34 30 90.9 676 2 G96593
35 30 90.9 978 2 T22681
36 30 90.9 1030 2 T18669
37 30 90.9 46 2 S10866
38 29 87.9 107 2 H84839
39 29 87.9 239 2 D85073
40 29 87.9 245 2 G81297
41 29 87.9 260 2 T17638
42 29 87.9 452 2 A12071
43 29 87.9 591 2 A69159
44 29 87.9 111C 1 B42544
45 29 87.9 1297 2 T39287

ALIGNMENTS

RESULT 1

C82247

probable proteinase inhibitor VC1057 [Imported] - Vibrio cholerae (strain N16961 ser
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82247
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller-
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: C82247

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <HPI>

A:Cross-references: GB:AE004187; GB:AE003852; NID:99655516; PIDN:AAF94216.1; GSPDB:G

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1057

A:Map position: 1

Query Match 100.0%; Score 33; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 4 CRLGC 8

RESULT 2

S75735

probable malate dehydrogenase (EC 1.1.1.37) - Synechocystis sp. (strain PCC 6803)

N:Alternate names: 2-ketoacid dehydrogenase; protein sll0891

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S75735

R:kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ya;

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75735

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-324 <KAN>

A:Cross-references: EMBL:D640C3; GB:AB001339; NID:gl001200; PIDN:BAA10470.1; PID:gl10C

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: cith
C:Superfamily: L-lactate dehydrogenase
C:Keywords: oxidoreductase

Query Match 100.0%; Score 33; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 280 CRLGC 284

RESULT 3

D84901
hypothetical protein At2g46320 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84901

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84901

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <STO>

A:Cross-references: GB:AE002093; MID:96598574; PIDN:AAF18629.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g46320

A:Map position: 2

Query Match 100.0%; Score 33; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
- Db 62 CRLGC 66

RESULT 4

B83963
hypothetical protein BH2506 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: B83963

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B83963

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <STO>

A:Cross-references: GB:AP001515; GB:BA000004; MID:g10174886; PIDN:BAB06225.1; GSPDB:GN00
A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2506

C:Superfamily: conserved hypothetical protein HI0365

Query Match 100.0%; Score 33; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 125 CRLGC 129

RESULT 5

AE3094

oxidoreductase Atu4377 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AE3094

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; MCCI
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AE3094

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-561 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL45171.1; PID:g17742847; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu4377

A:Map position: linear chromosome

Query Match 100.0%; Score 33; DB 2; Length 561;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 228 CRLGC 232

RESULT 6

E98192

probable oxidoreductase chain CJ0415 [imported] - Agrobacterium tumefaciens (strain C
C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002

C:Accession: E98192

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: E98192

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-561 <KUR>

A:Cross-references: GB:AE00781; PIDN:AAK89063.1; PID:g15158859; GSPDB:GN00170
A:Genetics:

A:Gene: AGR_L_980

A:Map position: linear chromosome

Query Match 100.0%; Score 33; DB 2; Length 561;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
|||||
Db 228 CRLGC 232

RESULT 7

AH2226

hypothetical protein all3367 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AH2226

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriqi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2226
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-783 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA075066.1; PID:g17132462; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all3367

Query Match 100.0%; Score 33; DB 2; Length 783;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 367 CRLGC 371

RESULT 8

AF2235
 hypothetical protein alr3437 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. strain PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AF2235
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF2235
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-791 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA075136.1; PID:g17132570; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3437

Query Match 100.0%; Score 33; DB 2; Length 791;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 364 CRLGC 368

RESULT 9

JQ1928
 G2-G1 polyprotein precursor - tomato spotted wilt virus (strain BR-01)
 N:Contains: glycoprotein G1; glycoprotein G2
 C:Species: tomato spotted wilt virus
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
 C:Accession: JQ1928
 R:Kornelink, R.; de Haan, P.; Meurs, C.; Peters, D.; Goldbach, R.
 J. Gen. Virol. 73, 2795-2804, 1992
 A:Title: The nucleotide sequence of the M RNA segment of tomato spotted wilt virus, a bunyavirus
 A:Reference number: JQ1927; MUID:93057351; PMID:1431808
 A:Accession: JQ1928
 A:Molecule type: genomic RNA
 A:Residues: 1-1135 <KOR>
 A:Cross-references: GB:S48091; NID:g259518; PIDN:AA024089.1; PID:g259520
 A:Note: the cleavage site between the G2 glycoprotein and G1 glycoprotein is not determined
 C:Genetics:
 A:Map position: segment M
 C:Superfamily: Impatiens necrotic spot virus G2-G1 polyprotein
 C:Keywords: glycoprotein; polyprotein; transmembrane protein
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-1135/Product: G2-G1 polyprotein #status predicted <GGP>
 F:315-335/Region: hydrophobic
 F:352-368/Region: hydrophobic
 F:448-483/Region: hydrophobic

F:1068-1089/Region: hydrophobic
 F:116,210,318,340,395,605,980,1095/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 100.0%; Score 33; DB 1; Length 1135;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 539 CRLGC 543

RESULT 10

G90296
 hypothetical protein SSO1395 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: G90296
 R:She, Q.; Jeffries, R.K.; Confalonieri, F.; Zivanovic, Y.; Aillard, G.; Awayez, M.J.; Redder, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: G90296
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1114 <KUR>
 A:Cross-references: GB:AE006641; NID:g13814614; PIDN:AAK41630.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SSO1395

Query Match 93.9%; Score 31; DB 2; Length 114;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 3 CRIGC 7

RESULT 11

AB2412
 hypothetical protein VCA0817 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: AB2412
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: AB2035; MUID:20406833; PMID:10952301
 A:Accession: AB2412
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-194 <HEI>
 A:Cross-references: GB:AE004410; GB:AE003853; NID:g9658244; PIDN:AAF96715.1; GSPDB:GN00155
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0817
 A:Map position: 2

Query Match 93.9%; Score 31; DB 2; Length 194;
 Best Local Similarity 80.0%; Pred. No. 82;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 |||||
 Db 122 CRIGC 126

RESULT 12

S71452
 trypsin/chymotrypsin inhibitor precursor - giant taro
 C:Species: Alocasia macrorrhiza (giant taro)
 C:Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 20-Jun-2000
 C:Accession: S71452; S42753
 R:Mathews, A.; Llewellyn, D.J.; Wu, Y.; Dennis, E.S.
 Plant Mol. Biol. 30, 1035-1039, 1996
 A:Title: Isolation and characterisation of full-length cDNA clones of the giant taro (Alocasia macrorrhiza) trypsin/chymotrypsin inhibitor from giant taro (Alocasia macrorrhiza)
 A:Reference number: S71452; MUID:96270380; PMID:8639740
 A:Accession: S71452
 A:Molecule type: mRNA
 A:Residues: 1-207 <MA>
 A:Cross-references: EMBL:X79606; NID:g1536883; PIDN:CAA56115.1; PID:g1536884
 A:Experimental source: corm
 R:Argall, M.E.; Bradbury, J.H.; Shaw, D.C.
 Biochim. Biophys. Acta 1204, 189-194, 1994
 A:Title: Amino-acid sequence of a trypsin/chymotrypsin inhibitor from giant taro (Alocasia macrorrhiza)
 A:Reference number: S42753; MUID:94191001; PMID:8142459
 A:Accession: S42753
 A:Molecule type: protein
 A:Residues: 19-67, 'E', 69-202 <ARG>
 C:Superfamily: plant Kunitz-type proteinase inhibitor
 C:Keywords: homodimer; polymorphism; proteinase inhibitor
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-203/Product: trypsin/chymotrypsin inhibitor #status experimental <MAT>
 F:204-207/Domain: carboxyl-terminal extension peptide #status predicted <CPRO>

Query Match 93.9%; Score 31; DB 2; Length 207;
 Best Local Similarity 80.0%; Pred. No. 85;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 II:II
 Db 165 CRMGC 169

RESULT 13
 A97113
 probable Fe-S-cluster redox enzyme, YLON B. subtilis ortholog [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
 C:Accession: A97113
 R:Nolling, J.J.; Brennon, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: A97113
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-345 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK79692.1; PID:g15024693; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC1726
 C:Superfamily: conserved hypothetical protein HI0365

Query Match 93.9%; Score 31; DB 2; Length 345;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 II:II
 Db 110 CRMGC 114

RESULT 14
 F69878
 conserved hypothetical protein ylon - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: F69878
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: F69878
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-363 <KUN>
 A:Cross-references: GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13448.1; PID:g26339 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: Ylon
 C:Superfamily: conserved hypothetical protein HI0365

Query Match 93.9%; Score 31; DB 2; Length 363;
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 II:II
 Db 126 CRIGC 130

RESULT 15
 E89894
 conserved hypothetical protein SA1061 [imported] - Staphylococcus aureus (strain N315
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: E89894
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K. C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: E89894
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <KUR>
 A:Cross-references: GB:BA000018; PID:g13701018; PIDN:BAB42313.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA1061
 C:Superfamily: conserved hypothetical protein HI0365

Query Match 93.9%; Score 31; DB 2; Length 364;
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 II:II
 Db 127 CRIGC 131

Search completed: August 5, 2003, 09:41:26
 Job time : 4.71429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:37:00 ; Search time 2 Seconds
(without alignments)
117.567 Million cell updates/sec

Title: US-09-883-727A-1_COPY_64_68
Perfect score: 33
Sequence: 1 CRLGC 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length	Description
1	33	100.0	82	1 YA57_VIBCH
2	33	100.0	313	1 TBLX_MOUSE
3	33	100.0	522	1 TBLX_HUMAN
4	33	100.0	526	1 TBLX_HUMAN
5	33	100.0	1135	1 VGLM_TSWV1
6	31	93.9	56	1 AMCL_APIME
7	31	93.9	184	1 ITC_ALOMA
8	31	93.9	348	1 LPXD_BARHE
9	31	93.9	363	1 YLON_BACSU
10	31	93.9	610	1 CALG_HUMAN
11	31	93.9	1357	1 YJ03_YEAST
12	30	90.9	303	1 MOVF_BMV
13	30	90.9	426	1 NUOF_AQUAE
14	30	90.9	586	1 SYD_EUCAI
15	30	90.9	1094	1 DPOL_GPCMV
16	29	87.9	118	1 GTHI_TRITC
17	29	87.9	971	1 RECK_HUMAN
18	29	87.9	971	1 RECK_MOUSE
19	29	87.9	1110	1 VGLM_INSV
20	29	87.9	2109	1 PKSL_ASPPA
21	28	84.8	84	1 YF7A_PSEAE
22	28	84.8	107	1 TAP1_ANTMA
23	28	84.8	109	1 GRP_TOBAC
24	28	84.8	109	1 NO16_SOYBN
25	28	84.8	123	1 WFD2_RABIT
26	28	84.8	126	1 THN1_WHEAT
27	28	84.8	127	1 THNA_HORVU
28	28	84.8	136	1 THN2_WHEAT
29	28	84.8	136	1 THNE_HORVU
30	28	84.8	136	1 THNE_WHEAT
31	28	84.8	137	1 GTH1_CORAU
32	28	84.8	137	1 GTH1_ONCKE
33	28	84.8	137	1 GTH1_ONCMA

AC Q90XE7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transducin beta-like 1X protein (Fragment).
 GN TBL1X OR TBL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RA Botcherby M.R.M., Straw R., Clarke D., Greyststrong J.S., Weston P.,
 RA Hunter G., Kimberly C., Rhodes M.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains at least 5 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; F38006; CAB61534.1; --
 DR MGD; MGI:1336172; Tbl1x.
 DR InterPro: IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR ProDom; PD000018; WD40; 2.
 DR SMART; SM00320; WD40; 5.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS50082; WD_REPEATS_2; 4.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT REPEAT 1 1 24 WD 1.
 FT REPEAT <1 24 WD 2.
 FT REPEAT 76 115 WD 3.
 FT REPEAT 160 198 WD 4.
 FT REPEAT 201 249 WD 5.
 FT REPEAT 252 291 WD 6.
 FT NON_TER 313 313
 SQ SEQUENCE 313 AA; 34808 MW; 3BC48683432DFEF7 CRC64;
 Query Match 100.0%; Score 33; DB 1; Length 313;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 147 CRLGC 151
 RESULT 3
 TBLX_HUMAN STANDARD; PRT; 522 AA.
 AC Q98087;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transducin beta-like 1Y protein (Transducin-beta-like 1, Y-linked).
 GN TBL1Y OR TBL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kuroda-Kawaguchi T., Skaletsky H., Minx P.J., Brown L.G., Rozen S.,
 RA Wilson R.K., Waterston R.H., Page D.C.;
 RT "The DNA sequence of the human Y chromosome.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 Lish domain.
 CC -1- SIMILARITY: Contains 8 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF332220; AAK13472.1; --
 DR EMBL; AF332221; AAK13473.1; --
 DR EMBL; AF332222; AAK13474.1; --
 DR Genew; HGNC:18502; TBL1Y.
 DR InterPro: IPR006594; Lish.
 DR ProDom; PD000018; WD40; 3.
 DR Pfam; PF00400; WD40; 8.
 DR SMART; SM00667; Lish; 1.
 DR SMART; SM00320; WD40; 8.
 DR PROSITE; PS50896; Lish; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS50082; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 4 36 LISH.
 FT REPEAT 177 216 WD 1.
 FT REPEAT 233 272 WD 2.
 FT REPEAT 274 313 WD 3.
 FT REPEAT 316 354 WD 4.
 FT REPEAT 357 396 WD 5.
 FT REPEAT 399 447 WD 6.
 FT REPEAT 450 489 WD 7.
 FT REPEAT 491 521 WD 8.
 SQ SEQUENCE 522 AA; 56688 MW; 4E02016422442D8 CRC64;
 Query Match 100.0%; Score 33; DB 1; Length 522;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 344 CRLGC 348
 RESULT 4
 TBLX_HUMAN STANDARD; PRT; 526 AA.
 AC O60907;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transducin beta-like 1X protein (Transducin-beta-like 1, X-linked).
 GN TBL1X OR TBL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99264241; PubMed=10330347;
 RA Bassi M.T., Ramesar R.S., Caciotti B., Winship I.M., De Grandi A.,
 RA Riboni M., Townes P.L., Beighton P., Ballabio A., Borsani G.;
 RT "X-linked late-onset sensorineural deafness caused by a deletion
 RT involving OAI and a novel gene containing WD-40 repeats.";
 RL Am. J. Hum. Genet. 64:1604-1616(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388957; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinacci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Contains 1 Lish domain.
CC -!- SIMILARITY: Contains 8 WD repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y12781; CAAT3319.1; ALT_INIT.
CC EMBL; BC032708; AAH32708.1; -.
CC Genew; HGNC:11585; TBLX.
CC MIM: 300196; -.
CC GO; GO:0005065; F-heterotrimeric G-protein; TAS.
CC GO; GO:0007605; P-hearing; TAS.
CC GO; GO:0007165; P-signal transduction; TAS.
CC GO; GO:0007601; P-vision; TAS.
CC InterPro; IPR006594; Lish.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 8.
CC PRINTS; PR00320; GPROTEINBRPT.
CC ProDom; PD000018; WD40; 3.
CC SMART; SM00667; Lish; 1.
CC SMART; SM00320; WD40; 1.
CC PROSITE; PS50896; Lish; 1.
CC PROSITE; PS00678; WD_REPEATS_1; 4.
CC PROSITE; PS50082; WD_REPEATS_2; 6.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 4 36 LISH.
FT REPEAT 179 218 WD 1.
FT REPEAT 235 274 WD 2.
FT REPEAT 276 315 WD 3.
FT REPEAT 318 358 WD 4.
FT REPEAT 359 398 WD 5.
FT REPEAT 401 449 WD 6.
FT REPEAT 452 491 WD 7.
FT REPEAT 493 525 WD 8.
FT DOMAIN 108 124 POLY-ALA.
SQ SEQUENCE 526 AA; 57048 MW; 98922F88EC42F6E9 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 346 CRLGC 350

RESULT 5

VGLM TSWV1
ID VGLM TSWV1 STANDARD; PRT; 1135 AA.
AC P36291.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE M polyprotein precursor [Contains: Glycoprotein G1; Glycoprotein G2].
GN M.
OS Tomato spotted wilt virus (strain Brazilian Br-01) (TSWV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
OX NCBI_TaxID:36413;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93057351; PubMed=1431808;
RA Kornelink R., de Haan P., Meurs C., Peters D., Goldbach R.;
RT "The nucleotide sequence of the M RNA segment of tomato spotted wilt
virus, a bunyavirus with two ambisense RNA segments.";
RL J. Gen. Virol. 73:2795-2804(1992).
RN [2]
RP ERRATUM.
RX MEDLINE=93224906; PubMed=8468562;
RA Kornelink R., de Haan P., Meurs C., Peters D., Goldbach R.;
RL J. Gen. Virol. 74:790-790(1993).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Probable).
CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC INCLUDING GLYCOPROTEIN G1 AND GLYCOPROTEIN G2.
CC -!- SIMILARITY: BELONGS TO THE HANTAVIRUSES M POLYPROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S48091; AAB24089.1; -.
CC PIR; J01928; J01928.
CC InterPro; IPR005167; Bunya_G1.
CC Pfam; PF03557; Bunya_G1; 1.
KW Polyprotein; Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1135 M POLYPROTEIN.
FT CHAIN 36 1135 GLYCOPROTEIN G1.
FT CHAIN 36 1135 GLYCOPROTEIN G2.
FT SITE 41 43 CELL ATTACHMENT SITE (POTENTIAL).
FT TRANSMEM 315 335 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 437 457 POTENTIAL.
FT TRANSMEM 749 769 POTENTIAL.
FT TRANSMEM 1068 1088 POTENTIAL.
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1095 1095 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1135 AA; 127317 MW; AC1A3FFFE84044FB CRC64;

Query Match 100.0%; Score 33; DB 1; Length 1135;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 539 CRLGC 543

RESULT 6

AMCI_APIME STANDARD; PRT; 56 AA.
ID AMCI_APIME

P56682;
 15-JUL-1999 (Rel. 38, Created)
 15-JUL-1999 (Rel. 38, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 Chymotrypsin inhibitor (AMCI).
 Oe Apis mellifera (Honeybee).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
 CC Apidae; Apis
 OX NCBI_TaxID=7460;
 RN [1]
 RP SEQUENCE, AND STRUCTURE BY NMR.
 RC TISSUE=Hemolymph;
 RA MEDLINE=9333935; PubMed=10411628;
 RX Bania J., Stachowiak D., Polanowski A.;
 RT "Primary structure and properties of the cathepsin G/chymotrypsin
 RL inhibitor from the larval hemolymph of Apis mellifera.";
 RL Eur. J. Biochem. 262:680-687(1999).
 CC -|- FUNCTION: Chymotrypsin and cathepsin G inhibitor.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
 DR PDB: 1CCV; 12-MAR-99.
 DR InterPro: IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL; 1.
 KW Serine protease inhibitor; 3D-structure.
 FT DOMAIN 3 56
 FT DISULFID 3 36
 FT DISULFID 12 32
 FT DISULFID 16 28
 FT DISULFID 20 56
 FT DISULFID 38 50
 FT TURN 5 6
 FT STRAND 7 13
 FT TURN 15 16
 FT STRAND 33 38
 FT TURN 40 41
 FT STRAND 43 45
 FT TURN 46 48
 FT STRAND 49 51
 FT HELIX 53 56
 SQ SEQUENCE 56 AA; 5973 MW; 092B2815AE6B2B7F CRC64;
 Query Match 93.9%; Score 31; DB 1; Length 56;
 Best Local Similarity 80.0%; Pred. No. 10;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 DB 32 CRIGC 36
 DB [1:11]
 RESULT 7
 ITC_ALOMA
 ID ITC_ALOMA STANDARD; PRT; 184 AA.
 AC P35812;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin/chymotrypsin inhibitor.
 OS Alocasia macrorrhiza (Giant taro).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Alocasia.
 OX NCBI_TaxID=4456;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. FUI-1;
 RX MEDLINE=94191001; PubMed=8142459;
 RA Argall M.E., Brandbury H.J., Shaw D.C.;
 RT "Amino-acid sequence of a trypsin/chymotrypsin inhibitor from giant
 RL taro (Alocasia macrorrhiza)."
 RL Biochim. Biophys. Acta 1204:189-194(1994).
 CC -|- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.
 CC -|- SUBUNIT: Homodimer.

-|- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
 FAMILY.
 CC HSP; P07596; IAVA.
 DR InterPro: IPR002160; Kunitz_legume.
 DR Pfam; PF00197; Kunitz_legume; 1.
 DR ProDom; PD000891; Kunitz_legume; 1.
 DR SMART; SM00452; STI; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; FALSE_NEG.
 KW Serine protease inhibitor.
 FT DISULFID 39 84
 FT DISULFID 136 147
 FT ACT_SITE 56 57
 FT VARIANT 24 24 M -> A (IN 50% OF THE CHAINS).
 FT VARIANT 50 50 E -> K (IN 25% OF THE CHAINS).
 SQ SEQUENCE 184 AA; 19774 MW; 60536E9C15F472ED CRC64;
 Query Match 93.9%; Score 31; DB 1; Length 184;
 Best Local Similarity 80.0%; Pred. No. 27;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 DB 147 CRMGC 151
 DB [1:11]
 RESULT 8
 LPXD_BARHE
 ID LPXD_BARHE STANDARD; PRT; 348 AA.
 AC Q8VQ23;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-3-O-[3-hydroxy-myristoyl] glucosamine N-acyltransferase
 DE (EC 2.3.1.-).
 GN LPXD.
 OS Bartonella henselae (Rochalimaea henselae).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Bartonellaceae; Bartonella.
 OX NCBI_TaxID=38323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zimmermann R., Augustin K., Schaal K., Sander A.;
 RT "Cloning, nucleotide sequencing, and expression of a hemin-binding
 RT protein of Bartonella henselae.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: UDP-3-O-(3-hydroxytetradecanoyl)glucosamine +
 CC (R)-3-hydroxytetradecanoyl-[acyl-carrier protein] = UDP-2,3-bis(3-
 CC hydroxytetradecanoyl)glucosamine + [acyl-carrier protein].
 CC -|- PATHWAY: Lipid A biosynthesis; third step.
 CC -|- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.
 CC LPXD SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF461795; AAL66375.1; -
 DR HAMAP; MF_00523; -; 1.
 DR InterPro: IPR001451; Hexapep_transf.
 DR Pfam; PF00132; hexapep; 9.
 DR Pfam; PF04613; Lpxd; 1.
 DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
 KW Transferase; Acyltransferase; Lipid A biosynthesis; Lipid synthesis;
 KW Repeat.
 SQ SEQUENCE 348 AA; 36836 MW; 7E8C57BA9ED50749 CRC64;
 Query Match 93.9%; Score 31; DB 1; Length 348;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
11:11
Db 165 CRIGC 169

RESULT 9

YLON_BACSU STANDARD; PRT; 363 AA.
AC Q34617:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ylon.
GN YLON.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;

[1]

SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98195738; PubMed=9534248;

RA Foulger D., Errington J.;

"A 28 kbp segment from the spoVM region of the Bacillus subtilis 168 genome.";

RL Microbiology 144:801-805(1998).

[2]

RN SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Boriuss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Broutillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,

RA Denton K., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,

RA Entian F., Errington J., Fabret C., Ferrich S., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;

"The complete genome sequence of the Gram-positive bacterium Bacillus

subtilis.";

RT Nature 390:249-255(1997).

[1]

RL Nature 390:249-255(1997).

CC -!- SIMILARITY: BELONGS TO THE UPF0063 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

[2]

CC EMBL; Y13937; CAA74265.1; -

CC EMBL; 299112; CAB13448.1; -

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR PIR; F69878; F69878.
DR Subtilist; BG13389; ylon.
DR InterPro; IPR004383; Cons_hypoth48.
DR InterPro; IPR006638; Elp3.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR00048; TIGR00048; 1.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 363 AA; 41569 MW; 628310764145C1C4 CRC64;

Query Match 93.9%; Score 31; DB 1; Length 363;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

Db 126 CRIGC 130

RESULT 10

CALG_HUMAN

ID CALG_HUMAN STANDARD; PRT; 610 AA.

AC O14967;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Calmegin precursor.

GN CLGN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=98094268; PubMed=9434179;

RA Tanaka H., Ikawa M., Tsuchida J., Nozaki M., Fujiwara T.,

RA Okabe M., Nishimune Y.;

"Cloning and characterization of the human Calmegin gene encoding

putative testis-specific chaperone.";

RL Gene 204:159-163(1997).

[2]

SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Bask S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Rosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

"Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: PROBABLY PLAYS AN IMPORTANT ROLE IN SPERMATOGENESIS.

CC BINDS CALCIUM IONS.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic

CC reticulum.

CC -!- TISSUE SPECIFICITY: Testis.

CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D86322; BAA22590.1; -;
 DR EMBL; BC028357; AAH28357.1; -;
 DR Genew; HGNC:2060; CLGN.
 DR MIM; 601858;
 DR GO; GO:0005783; C:Endoplasmic reticulum; TAS.
 DR GO; GO:0003754; F:chaperone activity; TAS.
 DR GO; GO:0007338; P:fertilization (sensu Animalia); TAS.
 DR InterPro; IPR001580; Calreticulin.
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PRODOM; PD001866; Calreticulin; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 DR Calcium-binding; Endoplasmic reticulum; Transmembrane; Signal;
 KW Repeat; Antigen.
 FT SIGNAL 1
 FT CHAIN 20
 FT DOMAIN 20 610
 FT TRANSMEM 472 492
 FT DOMAIN 493 610
 FT REPEAT 267 280
 FT REPEAT 284 297
 FT REPEAT 303 316
 FT REPEAT 322 335
 FT DOMAIN 339 397
 FT REPEAT 339 352
 FT REPEAT 356 369
 FT REPEAT 370 383
 FT REPEAT 384 397
 SQ SEQUENCE 610 AA; 70038 MW; F024FC4010D42D7E CRC64;

Query Match 93.9%; Score 31; DB 1; Length 610;
 Best Local Similarity 80.0%; Pred. No. 73;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CRLGC 5
 Db 351 CRIGC 355

RESULT 11
 YJ03_YEAST
 ID YJ03_YEAST STANDARD; PRT; 1357 AA.
 AC P47104;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 154.9 kDa protein in CPR7-PET191 intergenic region.
 GN YJ0333C OR J1590.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 1-1016 FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=95397595; PubMed=7668047;
 RA Huang M.-E., Chuat J.-C., Galibert F.;
 FT "Analysis of a 42.5 kb DNA sequence of chromosome X reveals three
 RT tRNA genes and 14 new open reading frames including a gene most
 RL probably belonging to the family of ubiquitin-protein ligases.";
 RL Yeast 11:775-781(1995).
 RN [2]
 RP SEQUENCE OF 728-1357 FROM N.A.
 RX MEDLINE=96109930; PubMed=8619316;

RA Zagulski M., Babinska B., Gromadka R., Migdalski A., Rytka J.,
 RA Sulicka J., Herbert C.J.;
 FT "The sequence of 24.3 kb from chromosome X reveals five complete open
 RT reading frames, all of which correspond to new genes, and a tandem
 RL insertion of a Ty1 transposon.";
 RL Yeast 11:1179-1186(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z49533; CAA89560.1; -;
 DR EMBL; L36344; AAA88735.1; -;
 DR EMBL; X87297; CAA60726.1; -;
 DR PIR; S57052; S57052.
 DR SGD; S0003794; RAV1.
 DR GO; GO:0005737; C:Cytoplasm; IDA.
 DR GO; GO:0000221; C:hydrogen-transporting ATPase VI domain; IPI.
 DR GO; GO:0007035; P:vacuolar acidification; IGI.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 3.
 DR SMART; SM00320; WD40; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 1357 AA; 154932 MW; 6B7270B6901274CF CRC64;

Query Match 93.9%; Score 31; DB 1; Length 1357;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CRLGC 5
 Db 1180 CRMGC 1184

RESULT 12
 MOV_P_BMV
 ID MOV_P_BMV STANDARD; PRT; 303 AA.
 AC P03603;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Cell-to-cell movement protein (MP) (3A protein).
 GN 3A.
 OS Brome mosaic virus (BMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Bromovirus.
 OX NCBI_TaxID=12302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82170484; PubMed=7338913;
 RA Ahlquist P., Luckow V., Kaesberg P.;
 RT "Complete nucleotide sequence of brome mosaic virus RNA3.";
 RL J. Mol. Biol. 153:23-38(1981).
 CC -1- FUNCTION: INVOLVED IN TRANSPORT OF THE VIRUS FROM THE INITIALLY
 CC INFECTED CELLS TO ADJACENT CELLS.
 CC -1- SIMILARITY: BELONGS TO THE CUCUMOVIRUSES/BROMOVIRUSES 3A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J02042; AAA46333.1; -;
 DR EMBL; V00099; CAA23434.1; -;
 DR PIR; A04214; P3BVAM.
 DR InterPro; IPR000603; 3A_mov.

DR Pfam; PF00803; 3A; 1.
 KW DNA-binding; Transport.
 SQ SEQUENCE 303 AA; 32482 MW; F66B93BC58372D37 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 303;
 Best Local Similarity 80.0%; Pred. No. 63;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 DB 47 CKLGC 51

RESULT 13

NUOF_AQUAE
 ID NUOF_AQUAE STANDARD; PRT; 426 AA.
 AC O66841;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NADH-quinone oxidoreductase chain F (EC 1.6.99.5) (NADH dehydrogenase
 I, chain F) (NDH-1, chain F).
 GN NUOF OR AQ_573.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaastland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
 sulfur (Fe-S) centers, to quinones in the respiratory chain.
 CC Couples the redox reaction to proton translocation (for every two
 CC electrons transferred, four hydrogen ions are translocated across
 CC the cytoplasmic membrane), and thus conserves the redox energy in
 CC a proton gradient (By similarity).
 CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
 CC -!- COFACTOR: Binds 1 FMN and 1 4Fe-4S cluster (Potential).
 CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 51 kDa SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000696; AAC06800.1; -.
 CC PIR; E70351; E70351.
 CC InterPro; IPR001949; Complex1_51k.
 CC Pfam; PF01512; Complex1_51k; 1.
 CC PROSITE; PS00644; Complex1_51k_1; 1.
 CC PROSITE; PS00645; Complex1_51k_2; 1.
 CC Oxidoreductase; NAD; Quinone; Flavoprotein; FMN; Iron-sulfur;
 KW 4Fe-4S; Complete proteome.
 FT NP_BIND 65 74 NAD (BY SIMILARITY).
 FT NP_BIND 176 223 FMN (BY SIMILARITY).
 FT NP_BIND 347 347 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 350 350 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 353 353 IRON-SULFUR (4FE-4S) (POTENTIAL).
 FT METAL 393 393 IRON-SULFUR (4FE-4S) (POTENTIAL).
 SQ SEQUENCE 426 AA; 47508 MW; A876712CBFA2E122 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 426;
 Best Local Similarity 80.0%; Pred. No. 83;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 DB 353 CRVGC 357

RESULT 14

SYD_BUCAI
 ID SYD_BUCAI STANDARD; PRT; 586 AA.
 AC P57401;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 DE (ASPRS).
 GN ASPS OR BU316.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP001119; BAB13024.1; -.
 CC HAMAP; P21889; 1EOR.
 CC HAMAP; MF_00044; 1.
 CC InterPro; IPR004524; AspS_bact.
 CC InterPro; IPR004115; GAD_dom.
 CC InterPro; IPR004364; tRNA-synt_2.
 CC InterPro; IPR002312; tRNA-synt_2.
 CC InterPro; IPR004365; tRNA-anti.
 CC InterPro; IPR006195; tRNA_ligase_II.
 CC Pfam; PF02938; GAD; 1.
 CC Pfam; PF00152; tRNA-synt_2; 2.
 CC Pfam; PF01336; tRNA-anti; 1.
 CC PRINTS; PR01042; TRNASYNTHASP.
 CC TIGRfams; TIGR00459; asps_bact; 1.
 CC PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 586 AA; 68120 MW; D7CB1B50A31800FE CRC64;

Query Match 90.9%; Score 30; DB 1; Length 586;
 Best Local Similarity 80.0%; Pred. No. 11e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
 DB 357 CKLGC 361

RESULT 15

DPOL_GPCMV

```

ID DPOL_GPCMV STANDARD; PRT; 1094 AA.
AC Q69025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
OS Guinea pig cytomegalovirus (strain 22122 / ATCC VR682) (GPCMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=103920;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97201603; PubMed=9049389;
RA Schleiss M.R.;
RT "Sequence and transcriptional analysis of the guinea-pig
cytomegalovirus DNA polymerase gene.";
RL J. Gen. Virol. 76:1827-1833(1995).
CC -|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA)(N).
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L25706; AAA43832.1;
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBc; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
DR Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1094 AA; 123052 MW; 647D254317B2B7FF CRC64;

Query Match 90.9%; Score 30; DB 1; Length 1094;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 92 CRVGC 96

```

Search completed: August 5, 2003, 09:39:48
Job time : 3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:37:25 ; Search time 8.42857 Seconds
(without alignments)
153.082 Million cell updates/sec

Title: US-09-883-727A-1_COPY_64_68

Perfect score: 33

Sequence: 1 CRLGC 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	35	6 Q8MIZ4	Q8miz4 capra hircu
2	33	100.0	69	3 Q8TGL6	Q8tgl6 saccharomyc
3	33	100.0	83	16 Q8EXP7	Q8exp7 shewanella
4	33	100.0	122	10 Q8GZ76	Q8gz76 arabidopsis
5	33	100.0	201	12 Q41933	Q41933 murid herpe
6	33	100.0	313	11 Q9D4P1	Q9d4p1 mus musculu
7	33	100.0	324	16 Q55383	Q55383 synecocyst
8	33	100.0	358	10 Q8S8M6	Q8s8m6 arabidopsis
9	33	100.0	361	10 Q8RY40	Q8ry40 arabidopsis
10	33	100.0	362	16 Q8K9Y8	Q8k9y8 bacillus ha
11	33	100.0	379	16 Q8KG98	Q8kg98 chlorobium
12	33	100.0	527	11 Q8BYO4	Q8byo4 mus musculu
13	33	100.0	527	11 Q8BMM0	Q8bmm0 mus musculu
14	33	100.0	561	16 Q8U7S2	Q8u7s2 agrobacteri
15	33	100.0	674	4 Q9H6M4	Q9h6m4 homo sapien
16	33	100.0	783	16 Q8YRS6	Q8yrs6 anabaena sp

17	33	100.0	791	16 Q8YRK9	Q8yrk9 anabaena sp
18	33	100.0	812	5 Q9VIS7	Q9vis7 drosophila
19	33	100.0	812	5 Q8T3K6	Q8t3k6 drosophila
20	33	100.0	860	10 Q9FTR7	Q9ftr7 oryza sativ
21	33	100.0	938	4 Q9HN7	Q9hn7 homo sapien
22	33	100.0	1050	11 Q8CCL4	Q8cccl4 mus musculu
23	33	100.0	1129	11 Q8BQX1	Q8bqx1 mus musculu
24	33	100.0	1135	12 Q9IKB7	Q9ikb7 tomato spot
25	33	100.0	1135	12 Q9IKB5	Q9ikb5 tomato spot
26	33	100.0	1135	12 Q55647	Q55647 tomato spot
27	33	100.0	1139	4 Q8IWI9	Q8iwi9 homo sapien
28	33	100.0	1151	10 Q9FPQ3	Q9fpq3 populus tri
29	33	100.0	3006	11 Q9QXJ5	Q9qxj5 mus musculu
30	33	100.0	5388	5 Q9U1D0	Q9uld0 leishmania
31	31	93.9	114	17 Q97YD1	Q97yd1 sulfolobus
32	31	93.9	131	10 Q8W3H6	Q8wh36 oryza sativ
33	31	93.9	185	10 Q8H8T9	Q8h8t9 oryza sativ
34	31	93.9	188	6 Q9GK46	Q9gk46 galago cras
35	31	93.9	194	16 Q9KLC9	Q9klc9 vibrilo chol
36	31	93.9	207	10 Q96230	Q96230 alocasia ma
37	31	93.9	345	16 Q97IC4	Q97ic4 clostridium
38	31	93.9	347	16 Q8XJL6	Q8xjl6 clostridium
39	31	93.9	364	16 Q99UQ0	Q99uq0 staphylococ
40	31	93.9	364	16 Q8NX16	Q8nx16 staphylococ
41	31	93.9	364	16 Q8CSW0	Q8csw0 staphylococ
42	31	93.9	534	5 Q9VY50	Q9vy50 drosophila
43	31	93.9	1222	5 Q9VY54	Q9vy54 drosophila
44	31	93.9	1274	10 Q9LVQ5	Q9lvq5 arabidopsis
45	31	93.9	1394	5 Q8MST1	Q8mst1 drosophila

ALIGNMENTS

RESULT 1

ID Q8MIZ4 PRELIMINARY; PRT; 35 AA.
AC Q8MIZ4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Butyrophillin precursor (fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim K., Hwang S., Kim Y., Park W., Kim N., Lee C.;
RT "Goat butyrophillin promoter region and exon 1";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY036083; AAL62339.1; -;
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3831 MA; ACBBE1F54157D058 CRC64;

Query Match 100.0%; Score 33; DB 6; Length 35;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRLGC 5
Db 8 CRLGC 12

RESULT 2

ID Q8TGL6 PRELIMINARY; PRT; 69 AA.
AC Q8TGL6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 7.9 kDa protein.
 GN YOR108C-A.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21624570; PubMed=11753363;
 RA Kumar A., Harrison P.M., Cheung K.H., Lan N., Echols N., Bertone P.,
 RA Miller P., Gerstein M.B., Snyder M.;
 RT "An integrated approach for finding overlooked genes in yeast.";
 RL Nat. Biotechnol. 20:58-63(2002).
 DR EMBL; AF479976; AAL79289.1;
 KW Hypothetical protein.
 SQ SEQUENCE 69 AA; 7936 MW; 3C50BB2BAF7EA996 CRC64;
 Query Match 100.0%; Score 33; DB 3; Length 69;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 46 CRLGC 50
 RESULT 3
 Q8EKP7
 ID Q8EKP7 PRELIMINARY; PRT; 83 AA.
 AC Q8EKP7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Conserved hypothetical protein.
 GN S00044.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MR-1;
 RX MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.,
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 Shewanella oneidensis";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 DR EMBL; AE015456; AAN53131.1;
 DR TIGR; S00044;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 83 AA; 8721 MW; E53E67FF419D6B5 CRC64;
 Query Match 100.0%; Score 33; DB 16; Length 83;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 3 CRLGC 7
 RESULT 4
 Q8GZ76
 ID Q8GZ76 PRELIMINARY; PRT; 122 AA.
 AC Q8GZ76;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN AF2G46320.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia;
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RT "Arabidopsis thaliana full-length cDNA.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK117170; BAC41848.1;
 KW Hypothetical protein.
 SQ SEQUENCE 122 AA; 13173 MW; 780BCBDF0F0749D7 CRC64;
 Query Match 100.0%; Score 33; DB 10; Length 122;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRLGC 5
 Db 62 CRLGC 66
 RESULT 5
 Q41933
 ID Q41933 PRELIMINARY; PRT; 201 AA.
 AC Q41933;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE BHV4-1E1 homolog (CONTAINS potential zinc-binding motif and TWO
 DE hydrophobic domains).
 GN GAMMAHV.K3 OR 12.
 OS Murid herpesvirus 4.
 OC Viruses; GSDNA viruses, non-RTA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 OX NCBI_TaxID=33708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=WUMS, and G2.4;
 RX MEDLINE=97366649; PubMed=9223479;
 RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
 RA Dal Canto A.J., Speck S.H.;
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 68.";
 RL J. Virol. 71:5894-5904(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=WUMS;
 RA Latreille P., Wamsley P., Waterston R.H.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=G2.4;
 RA Milligan S., Efsthathiou S., Stewart J.P., Nash A.A., Davison A.J.;
 RT "Genetic content of murine gammaherpesviruses.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MHV76;
 RA Macrae A.I., Milligan S., Dutia B.M., Davison A.J., Nash A.A.,
 RA Stewart J.P.;
 RT "The left end of the murine herpesvirus 68 genome plays a critical
 RT role in pathogenesis.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; U97553; AAB66429.1; -.
DR EMBL; AF105037; AAF19280.1; -.
DR EMBL; AF324455; AAK16705.1; -.
DR InterPro; IPR001005; Myb_DNA_binding.
DR PROSITE; PS00037; MYB_1; 1.
SQ SEQUENCE 201 AA; 22214 MW; 0953B5BDACDF8367 CRC64;

Query Match 100.0%; Score 33; DB 12; Length 201;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 191 CRLGC 195

RESULT 6
Q9D4P1 PRELIMINARY; PRT; 313 AA.
AC Q9D4P1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 4930585L22Rik protein.
GN 4930585L22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK016353; BAB30202.1; -.
DR MGD; MGI:1923141; 4930585L22Rik.
DR InterPro; IPR000048; IQ_region.
DR Pfam; PF00612; IQ; 1.
DR SMART; SM00015; IQ; 1.
SQ SEQUENCE 313 AA; 37042 MW; A8FC20B1CC8AFB85 CRC64;

Query Match 100.0%; Score 33; DB 11; Length 313;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 307 CRLGC 311

RESULT 7
Q55383 PRELIMINARY; PRT; 324 AA.
ID Q55383
AC Q55383;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2-ketoacid dehydrogenase (Malate dehydrogenase, lactate
DE dehydrogenase).
DE CITH OR SLL0891.
GN Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PCC 6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugliura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugliura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D64003; BAA10470.1; -.
DR HSSP; Q27743; ICET.
DR InterPro; IPR001236; 1dh.
DR InterPro; IPR001557; LLDH.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00056; 1dh; 1.
DR Pfam; PF02866; 1dh_C; 1.
DR PRINTS; PR00086; LLDHGRNASE.
KW Complete proteome.
SQ SEQUENCE 324 AA; 34345 MW; 7A0B773CEF6528FF CRC64;

Query Match 100.0%; Score 33; DB 16; Length 324;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
DB 280 CRLGC 284

RESULT 8
Q8S8M6 PRELIMINARY; PRT; 358 AA.
ID Q8S8M6
AC Q8S8M6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 38.8 kDa protein.
GN AT2G46320.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

```

```

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005397; AAL15049.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein.
SQ SEQUENCE 358 AA; 38788 MW; F94659415D772601 CRC64;

Query Match 100.0%; Score 33; DB 10; Length 358;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 62 CRLGC 66

RESULT 9
Q8RY40 PRELIMINARY; PRT; 361 AA.
ID Q8RY40
AC Q8RY40
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE At2g46320/FilC10.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kwai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY078922; AAL84928.1; -
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
SQ SEQUENCE 361 AA; 39146 MW; F2F7EF20F8D6990A CRC64;

Query Match 100.0%; Score 33; DB 10; Length 361;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 62 CRLGC 66

RESULT 10
Q9K9Y8 PRELIMINARY; PRT; 362 AA.
ID Q9K9Y8
AC Q9K9Y8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein BH2506.
GN BH2506.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001515; BAB06225.1; -
DR InterPro; IPR004383; Cons_hypoth48.
DR InterPro; IPR006638; Elp3.
DR SMART; SM00729; Elp3; 1.
DR TIGRfams; TIGR00048; TIGR00048; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 362 AA; 41389 MW; AE14020C99DFCC31 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 362;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5
Db 125 CRLGC 129

RESULT 11
Q8KG98 PRELIMINARY; PRT; 379 AA.
ID Q8KG98
AC Q8KG98;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aminotransferase, class V.
GN Cr0070.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RL photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE012787; AAM71318.1; -
DR TIGR; CT0070; -
DR InterPro; IPR000192; Aminotransfv.
DR Pfam; PF00266; aminotran_5; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 379 AA; 41767 MW; FA7226544514DE75 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRLGC 5

```

```

Db          274 CRLGC 278
|||||
RESULT 12
Q8BYQ4      PRELIMINARY;      PRT;      527 AA.
AC Q8BYQ4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transducin.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK038674; BAC30092.1; -.
SQ SEQUENCE 527 AA; 56802 MW; 146435A9C51DFDA6 CRC64;

Query Match          100.0%; Score 33; DB 11; Length 527;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 CRLGC 5
|||||
Db          347 CRLGC 351

RESULT 13
Q8BNM0
ID Q8BNM0      PRELIMINARY;      PRT;      527 AA.
AC Q8BNM0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transducin.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=pituitary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK030547; BAC27015.1; -.
SQ SEQUENCE 527 AA; 56832 MW; 0466F764ABA25CB0 CRC64;

Query Match          100.0%; Score 33; DB 11; Length 527;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 CRLGC 5
|||||
Db          347 CRLGC 351

RESULT 14
Q8U7S2
ID Q8U7S2      PRELIMINARY;      PRT;      561 AA.

```

```

AC Q8U7S2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Oxidoreductase.
GN ATU4377 OR AGR_L_980.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RL Science 294:2317-2323(2001).";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lapps C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RL Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).";
DR EMBL; AE009366; AAL45171.1; -.
DR EMBL; AE008247; AAK89063.1; -.
KW Complete proteome.
SQ SEQUENCE 561 AA; 63350 MW; A4ADF1015052B94A CRC64;

Query Match          100.0%; Score 33; DB 16; Length 561;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 CRLGC 5
|||||
Db          228 CRLGC 232

RESULT 15
Q9H6M4
ID Q9H6M4      PRELIMINARY;      PRT;      674 AA.
AC Q9H6M4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ72098 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025751; BAB15232.1; -.
KW Hypothetical protein.

```

FT NON_TER 674 674
SQ SEQUENCE 674 AA; 75293 MW; 4A3BC3D73EA2D5FB CRC64;
Query Match 100.0%; Score 33; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRLGC 5
IIIII
Db 463 CRLGC 467

Search completed: August 5, 2003, 09:40:55
Job time : 10.4286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:36:20 ; Search time 62.5714 Seconds
(without alignments)
76.102 Million cell updates/sec

Title: US-09-883-727A-53
Perfect score: 185
Sequence: 1 GCDGFKRLGCTYGFRTDKKGCEAFCTCNT 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	185	100.0	30	AAE20000	Complement C1s cat
2	185	100.0	122	AAE19948	Haementaria ghilia
3	185	100.0	142	AA106449	Leech haemostasin.
4	185	100.0	142	AA106450	Leech haemostasin.
5	180	97.3	107	AA106451	Leech haemostasin
6	179	96.8	29	AAE20068	Complement C1s cat
7	175	94.6	30	AAE20004	Complement C1s cat
8	175	94.6	30	AAE20005	Complement C1s cat
9	175	94.6	30	AAE20006	Complement C1s cat

10	175	94.6	30	23	AAE20007	Complement C1s cat
11	175	94.6	30	23	AAE20018	Complement C1s cat
12	175	94.6	30	23	AAE20019	Complement C1s cat
13	175	94.6	30	23	AAE20020	Complement C1s cat
14	175	94.6	30	23	AAE20021	Complement C1s cat
15	175	94.6	30	23	AAE20032	Complement C1s cat
16	175	94.6	30	23	AAE20033	Complement C1s cat
17	175	94.6	30	23	AAE20034	Complement C1s cat
18	175	94.6	30	23	AAE20035	Complement C1s cat
19	165	89.2	30	23	-AAE20008	Complement C1s cat
20	165	89.2	30	23	AAE20009	Complement C1s cat
21	165	89.2	30	23	AAE20010	Complement C1s cat
22	165	89.2	30	23	AAE20011	Complement C1s cat
23	165	89.2	30	23	AAE20012	Complement C1s cat
24	165	89.2	30	23	AAE20013	Complement C1s cat
25	165	89.2	30	23	AAE20022	Complement C1s cat
26	165	89.2	30	23	AAE20023	Complement C1s cat
27	165	89.2	30	23	AAE20024	Complement C1s cat
28	165	89.2	30	23	AAE20025	Complement C1s cat
29	165	89.2	30	23	AAE20026	Complement C1s cat
30	165	89.2	30	23	AAE20027	Complement C1s cat
31	165	89.2	30	23	AAE20036	Complement C1s cat
32	165	89.2	30	23	AAE20037	Complement C1s cat
33	165	89.2	30	23	AAE20038	Complement C1s cat
34	165	89.2	30	23	AAE20039	Complement C1s cat
35	165	89.2	30	23	AAE20040	Complement C1s cat
36	165	89.2	30	23	AAE20041	Complement C1s cat
37	155	83.8	30	23	AAE20014	Complement C1s cat
38	155	83.8	30	23	AAE20015	Complement C1s cat
39	155	83.8	30	23	AAE20016	Complement C1s cat
40	155	83.8	30	23	AAE20017	Complement C1s cat
41	155	83.8	30	23	AAE20028	Complement C1s cat
42	155	83.8	30	23	AAE20029	Complement C1s cat
43	155	83.8	30	23	AAE20030	Complement C1s cat
44	155	83.8	30	23	AAE20031	Complement C1s cat
45	155	83.8	30	23	AAE20042	Complement C1s cat

ALIGNMENTS

RESULT 1
AAE20000
ID AAE20000 standard; peptide; 30 AA.

AC AAE20000;

XX 18-JUN-2002 (first entry)

DT Complement C1s catalytic site-directed moiety, peptide #1.

DE Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute reagent; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
restenosis; myasthenia gravis.

XX Unidentified.

OS WO200198365-A2.

PN 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

PF 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

PI West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX

PT New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis

XX Claim 7; Page 49; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejections, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anemias, or myasthenia gravis. The present
 CC peptide sequence is complement C1s catalytic site-directed moiety.

XX Sequence 30 AA;

Query Match 100.0%; Score 185; DB 23; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.1e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 30
 |||||
 Db 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 30

RESULT 2

AAE19948
 ID AAE19948 standard; Protein; 122 AA.

XX AAE19948;

XX 18-JUN-2002 (first entry)

DE Haementaria ghilianii BD001 protein.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis; BD001.

XX Haementaria ghilianii.

OS WO200198365-A2.

PN 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis

XX Claim 6; Page 2; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejections, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anemias, or myasthenia gravis. The present
 CC sequence is Haementaria ghilianii BD001 protein used in the invention.

XX Sequence 122 AA;

Query Match 100.0%; Score 185; DB 23; Length 122;
 Best Local Similarity 100.0%; Pred. No. 6.7e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 30
 |||||
 Db 58 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 87

RESULT 3

AAAY06449
 ID AAAY06449 standard; Protein; 142 AA.

XX AAAY06449;

XX 27-SEP-1999 (first entry)

DE Leech haemostasin.

XX Haemostasin; leech; blood clotting; coagulation; complement;
 KW Factor XII; C1 esterase; inhibitor; cardiovascular disease;
 KW inflammation; autoimmune disease; thrombosis; pulmonary embolus;
 KW angioplasty; endarterectomy; haemodialysis; cardiopulmonary bypass;
 KW organ rejection; sepsis; myocardial infarction; stroke;
 KW ischaemic shock; atherosclerosis; vasculitis; rheumatoid arthritis;
 KW sickle cell anaemia; angioedema; lupus arthritis; nephritis;
 KW glomerulonephritis; systemic sclerosis; Behcet's syndrome;
 KW cerebral lupus; Guillain-Barre disease; multiple sclerosis;
 KW myasthenia gravis; pemphigus; bullous pemphigoid; burn; asthma;
 KW phototoxic reaction; skin reaction; infection; thyroiditis;
 KW inflammatory bowel disease; Alzheimer's disease;
 KW paroxysmal nocturnal haemoglobinuria; haemolytic anaemia; therapy.

XX Haementaria ghilianii.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "leader peptide"

FT 21..142

FT /note= "mature protein, specifically claimed in

FT Claim 5"

XX WO936439-A1.

XX 22-JUL-1999.

XX 15-JAN-1999; 99WO-GB00147.

XX 16-JAN-1998; 98GB-0000817.

XX (BIOD-) BIO-DISCOVERY LTD.

XX Finney S, Seale L;

XX WPI; 1999-430600/36.

XX N-PSDB; AAX59408.

XX New polypeptides which inhibit C1 esterase and/or factor XII
 PT activation are used to treat cardiovascular, inflammatory or
 PT autoimmune disorders

XX Claim 5; Page 5-6; 43pp; English.

XX The present sequence represents a haemostasin polypeptide of
 CC the leech Haementaria ghilianii. The 14 kDa mature portion of the

polypeptide is claimed. It acts as an inhibitor of the initiation complexes of blood coagulation and/or of complement. 3 Variants (see also AAY06450 and AAY06451) of haemostasin were identified that vary at positions 16 and 60 of the mature protein. The invention provides vectors and host cells suitable for production of haemostasin. A claimed method for the treatment or prophylaxis of a condition or disorder related to Cl esterase initiation and/or Factor XII activation comprises the administration to a patient of an effective, inhibitory amount of a haemostasin polypeptide. The polypeptide may administered by means of being expressed in the cells of the patient, the cells comprising and expressing a haemolysin nucleic acid sequence. Haemostasins can potentially be used to inhibit the activation of coagulation in the treatment of thrombotic disease such as deep venous thrombosis, pulmonary embolus, and thrombosis associated with angioplasty and endarterectomy. Disease may also be alleviated by the ability of the haemostasins to inhibit both complement activation and the intrinsic pathway of blood coagulation, such as in haemodialysis, cardiopulmonary bypass, or rejection of transplanted organs or tissues, or in cases of sepsis, myocardial infarction, stroke, particularly in the injury caused to tissues by reperfusion after an ischaemic period, atherosclerosis, shock, vasculitis, rheumatoid arthritis, sickle cell anaemia or angioedema. The haemostasins may also be used in conditions associated with activation of complement, such as: various autoimmune diseases (e.g. lupus arthritis); glomerulonephritis; nephritis; nephropathy; systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillan-Barre disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous pemphigoid; phototoxic reactions; thermal burns; anaphylaxis; asthma; skin reactions; infections; inflammatory bowel disease; thyroiditis; infertility; Alzheimer's disease; paroxysmal nocturnal haemoglobinuria; and haemolytic anaemia (all claimed).

Sequence 142 AA;

Query Match 100.0%; Score 185; DB 20; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.6e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGDGFKRLGCTGYGFKDKGCEAFCTCNT 30
|||||
DB 78 GGDGFKRLGCTGYGFKDKGCEAFCTCNT 107

RESULT 4
AAY06450
ID AAY06450 standard; Protein; 142 AA.

AC AAY06450;

XX 27-SEP-1999 (first entry)

XX Leech haemostasin.

XX Haemostasin; leech; blood clotting; coagulation; complement;
KW Factor XII; Cl esterase; inhibitor; cardiovascular disease;
KW inflammation; autoimmune disease; thrombosis; pulmonary embolus;
KW angioplasty; endarterectomy; haemodialysis; cardiopulmonary bypass;
KW organ rejection; sepsis; myocardial infarction; stroke;
KW ischaemic shock; atherosclerosis; vasculitis; rheumatoid arthritis;
KW sickle cell anaemia; angioedema; lupus arthritis; nephritis;
KW glomerulonephritis; systemic sclerosis; Behcet's syndrome;
KW cerebral lupus; Guillan-Barre disease; multiple sclerosis;
KW myasthenia gravis; pemphigus; bullous pemphigoid; burn; asthma;
KW phototoxic reaction; skin reaction; infection; thyroiditis;
KW inflammatory bowel disease; Alzheimer's disease;
KW paroxysmal nocturnal haemoglobinuria; haemolytic anaemia; therapy.

XX Haementeria ghillanii.

XX Key Location/Qualifiers

FH Peptide 1..20

FT /note= "leader peptide"

FT Protein 21..142

FT /note= "mature protein, specifically claimed in Claim 5"

XX WO9936439-A1.

XX 22-JUL-1999.

XX 15-JAN-1999; 99WO-GB00147.

XX 16-JAN-1998; 98GB-0000817.

XX (BIOD-) BIO-DISCOVERY LTD.

XX Finney S, Seale L;

XX WPI; 1999-430600/36.

XX New polypeptides which inhibit Cl esterase and/or factor XII activation are used to treat cardiovascular, inflammatory or autoimmune disorders

XX Claim 5; Page 36; 43pp; English.

XX The present sequence represents a haemostasin polypeptide of the leech *Haementeria ghillanii*. The 14 kDa mature portion of the polypeptide is claimed. It acts as an inhibitor of the initiation complexes of blood coagulation and/or of complement. 3 Variants (see also AAY06449 and AAY06451) of haemostasin were identified that vary at positions 16 and 60 of the mature protein. The invention provides vectors and host cells suitable for production of haemostasin. A claimed method for the treatment or prophylaxis of a condition or disorder related to Cl esterase initiation and/or Factor XII activation comprises the administration to a patient of an effective, inhibitory amount of a haemostasin polypeptide. The polypeptide may administered by means of being expressed in the cells of the patient, the cells comprising and expressing a haemolysin nucleic acid sequence. Haemostasins can potentially be used to inhibit the activation of coagulation in the treatment of thrombotic disease such as deep venous thrombosis, pulmonary embolus, and thrombosis associated with angioplasty and endarterectomy. Disease may also be alleviated by the ability of the haemostasins to inhibit both complement activation and the intrinsic pathway of blood coagulation, such as in haemodialysis, cardiopulmonary bypass, or rejection of transplanted organs or tissues, or in cases of sepsis, myocardial infarction, stroke, particularly in the injury caused to tissues by reperfusion after an ischaemic period, atherosclerosis, shock, vasculitis, rheumatoid arthritis, sickle cell anaemia or angioedema. The haemostasins may also be used in conditions associated with activation of complement, such as: various autoimmune diseases (e.g. lupus arthritis); glomerulonephritis; nephritis; nephropathy; systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillan-Barre disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous pemphigoid; phototoxic reactions; thermal burns; anaphylaxis; asthma; skin reactions; infections; inflammatory bowel disease; thyroiditis; infertility; Alzheimer's disease; paroxysmal nocturnal haemoglobinuria; and haemolytic anaemia (all claimed).

XX Sequence 142 AA;

Query Match 100.0%; Score 185; DB 20; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.6e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGDGFKRLGCTGYGFKDKGCEAFCTCNT 30
|||||
DB 78 GGDGFKRLGCTGYGFKDKGCEAFCTCNT 107

RESULT 5

AAY06451

ID AAY06451 standard; Protein; 107 AA.

XX AAY06451;
 XX 27-SEP-1999 (first entry)
 XX Leech haemostasin (partial sequence).
 XX Haemostasin; leech; blood clotting; coagulation; complement;
 KW Factor XII; Cl esterase; inhibitor; cardiovascular disease;
 KW inflammation; autoimmune disease; thrombosis; pulmonary embolus;
 KW angioplasty; endarterectomy; haemodialysis; cardiopulmonary bypass;
 KW organ rejection; sepsis; myocardial infarction; stroke;
 KW ischaemic shock; atherosclerosis; vasculitis; rheumatoid arthritis;
 KW sickle cell anaemia; angioedema; lupus arthritis; nephritis;
 KW glomerulonephritis; systemic sclerosis; Behcet's syndrome;
 KW cerebral lupus; Guillan-Barre disease; multiple sclerosis;
 KW myasthenia gravis; pemphigus; bullous pemphigoid; burn; asthma;
 KW phototoxic reaction; skin reaction; infection; thyroiditis;
 KW inflammatory bowel disease; Alzheimer's disease;
 KW paroxysmal nocturnal haemoglobinuria; haemolytic anaemia; therapy.
 XX Haementeria ghilianii.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /note= "leader peptide"
 FT Protein 21..107
 FT /note= "mature protein, specifically claimed in
 FT Claim 5"
 XX
 XX WO9936439-A1.
 XX
 XX 22-JUL-1999.
 XX
 XX 15-JAN-1999; 99WO-GB00147.
 XX
 XX 16-JAN-1998; 98GB-0000817.
 XX
 XX (BIOD-) BIO-DISCOVERY LTD.
 XX
 XX Finney S, Seale L;
 XX
 XX WPT; 1999-430600/36.
 XX
 XX New polypeptides which inhibit Cl esterase and/or factor XII
 XX activation are used to treat cardiovascular, inflammatory or
 XX autoimmune disorders
 XX
 XX Claim 5; Page 36-37; 43pp; English.
 XX
 XX The present sequence represents a haemostasin polypeptide of
 XX the leech Haementeria ghilianii. The 14 kDa mature portion of the
 XX polypeptide is claimed. It acts as an inhibitor of the initiation
 XX complexes of blood coagulation and/or of complement. 3 variants
 XX (see also AAY06449 and AAY06450) of haemostasin were identified that
 XX vary at positions 16 and 60 of the mature protein. The invention
 XX provides vectors and host cells suitable for production of
 XX haemostasin. A claimed method for the treatment or prophylaxis of
 XX a condition or disorder related to Cl esterase initiation and/or
 XX Factor XII activation comprises the administration to a patient of
 XX an effective, inhibitory amount of a haemostasin polypeptide. The
 XX polypeptide may administered by means of being expressed in the
 XX cells of the patient, the cells comprising and expressing a
 XX haemolysin nucleic acid sequence. Haemostasins can potentially be
 XX used to inhibit the activation of coagulation in the treatment of
 XX thrombotic disease such as deep venous thrombosis, pulmonary
 XX embolus, and thrombosis associated with angioplasty and
 XX endarterectomy. Disease may also be alleviated by the ability of the
 XX haemostasins to inhibit both complement activation and the intrinsic
 XX pathway of blood coagulation, such as in haemodialysis,
 XX cardiopulmonary bypass, or rejection of transplanted organs or
 XX tissues, or in cases of sepsis, myocardial infarction, stroke,
 XX particularly in the injury caused to tissues by reperfusion after

CC an ischaemic period, atherosclerosis, shock, vasculitis,
 CC rheumatoid arthritis, sickle cell anaemia or angioedema. The
 CC haemostasins may also be used in conditions associated with
 CC activation of complement, such as: various autoimmune diseases
 CC (e.g. lupus arthritis); glomerulonephritis; nephritis; nephropathy;
 CC systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillan-Barre
 CC disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous
 CC pemphigoid; phototoxic reactions; thermal burns; anaphylaxis;
 CC asthma; skin reactions; infections; inflammatory bowel disease;
 CC thyroiditis; infertility; Alzheimer's disease; paroxysmal
 CC nocturnal haemoglobinuria; and haemolytic anaemia (all claimed).
 XX
 XX Sequence 107 AA;
 XX
 XX Query Match 97.3%; Score 180; DB 20; Length 107;
 XX Best Local Similarity 96.7%; Pred. No. 2.2e-13;
 XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 30
 DB ||:|||||:|||||:|||||:|||||:|||||
 78 GCNGFKRLGCTGYGFKTDKKGCEAFCTCNT 107
 RESULT 6
 AAE20068
 ID AAE20068 standard; peptide; 29 AA.
 XX
 XX AC AAE20068;
 XX
 XX 18-JUN-2002 (first entry)
 XX
 XX Complement Cls catalytic site-directed moiety, peptide #67.
 KW Complement Cls; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 XX Unidentified.
 XX
 XX WO200198365-A2.
 XX
 XX 27-DEC-2001.
 XX
 XX 18-JUN-2001; 2001WO-US19405.
 XX
 XX 21-JUN-2000; 2000US-212998P.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX West RR, Sheppard PO, Fox BA;
 XX
 XX WPI; 2002-241177/29.
 XX
 XX New complement Cls inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 XX infarct, stroke), hyperacute rejectin, sepsis, burns, wound healing,
 XX asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 XX serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 XX peptide sequence is complement Cls catalytic site-directed moiety.
 XX
 XX Claim 9; Page 50; 99pp; English.
 XX
 XX The invention relates to a polypeptide that inhibits complement Cls. The
 XX inhibitory peptides are useful as therapeutic agents, as preservatives
 XX in blood samples, and in affinity purification procedures to isolate Cls.
 XX Molecules that inhibit complement may be used for treating diseases
 XX in which complement activation has been shown to occur, e.g., adult
 XX respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 XX infarct, stroke), hyperacute rejectin, sepsis, burns, wound healing,
 XX asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 XX serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 XX peptide sequence is complement Cls catalytic site-directed moiety.

RESULTS

xx
de
Complement C1s catalytic site-directed moiety, peptide #1.

xx
Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
preservative; ischaemia-perfusion injury; myocardial infarct; sepsis;
hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;

KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX Unidentified.
 OS WO200198365-A2.
 XX
 PN
 XX
 PD 27-DEC-2001.
 XX
 XX 18-JUN-2001; 2001WO-US19405.
 PF
 XX 21-JUN-2000; 2000US-212998P.
 PR
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX West RR, Sheppard PO, Fox BA;
 XX WPI; 2002-241177/29.
 XX
 XX New complement C1s inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 XX sepsis
 XX
 PS Disclosure; Page 18; 99pp; English.
 XX
 XX The invention relates to a polypeptide that inhibits complement C1s. The
 XX inhibitory peptides are useful as therapeutic agents, as preservatives
 XX in blood samples, and in affinity purification procedures to isolate C1s.
 XX Molecules that inhibit complement may be used for treating diseases
 XX in which complement activation has been shown to occur, e.g., adult
 XX respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 XX infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 XX asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 XX serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 XX peptide sequence is complement C1s catalytic site-directed moiety.
 XX
 PS Sequence 30 AA;
 XX
 XX Query Match 94.6%; Score 175; DB 23; Length 30;
 XX Best Local Similarity 96.7%; Pred. No. 2.8e-13;
 XX Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 GCDGFKRLGCTGYGFKTKKGCEAFCTCNT 30
 |||||
 Db 1 GCDGFKRLGCTGYGFKTKKGCEAFCTCNT 30
 |||||
 RESULT 10
 AAE20007
 ID AAE20007 standard; peptide; 30 AA.
 AC AAE20007;
 XX
 XX 18-JUN-2002 (first entry)
 DT
 XX
 DE Complement C1s catalytic site-directed moiety, peptide #8.
 XX
 KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.
 XX
 XX WO200198365-A2.
 PN
 XX
 XX 27-DEC-2001.
 PD
 XX
 XX 18-JUN-2001; 2001WO-US19405.
 PF
 XX 21-JUN-2000; 2000US-212998P.
 PR

XX (ZYMO) ZYMOGENETICS INC.
 PA West RR, Sheppard PO, Fox BA;
 XX WPI; 2002-241177/29.
 XX
 XX New complement C1s inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 XX sepsis
 XX
 PS Disclosure; Page 18; 99pp; English.
 XX
 XX The invention relates to a polypeptide that inhibits complement C1s. The
 XX inhibitory peptides are useful as therapeutic agents, as preservatives
 XX in blood samples, and in affinity purification procedures to isolate C1s.
 XX Molecules that inhibit complement may be used for treating diseases
 XX in which complement activation has been shown to occur, e.g., adult
 XX respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 XX infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 XX asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 XX serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 XX peptide sequence is complement C1s catalytic site-directed moiety.
 XX
 PS Sequence 30 AA;
 XX
 XX Query Match 94.6%; Score 175; DB 23; Length 30;
 XX Best Local Similarity 96.7%; Pred. No. 2.8e-13;
 XX Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 GCDGFKRLGCTGYGFKTKKGCEAFCTCNT 30
 |||||
 Db 1 GCDGFKRLGCTGYGFKTKKGCEAFCTCNT 30
 |||||
 RESULT 11
 AAE20018
 ID AAE20018 standard; peptide; 30 AA.
 AC AAE20018;
 XX
 XX 18-JUN-2002 (first entry)
 DT
 XX
 DE Complement C1s catalytic site-directed moiety, peptide #19.
 XX
 KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.
 XX
 XX WO200198365-A2.
 PN
 XX
 XX 27-DEC-2001.
 PD
 XX
 XX 18-JUN-2001; 2001WO-US19405.
 PF
 XX 21-JUN-2000; 2000US-212998P.
 PR
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX West RR, Sheppard PO, Fox BA;
 XX WPI; 2002-241177/29.
 XX
 XX New complement C1s inhibitor polypeptides for treating diseases in
 XX which complement activation has been shown to occur, e.g. adult
 XX respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 XX sepsis
 XX

PS Disclosure; Page 18; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
XX inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
XX peptide sequence is complement C1s catalytic site-directed moiety.

XX Sequence 30 AA;

Query Match 94.6%; Score 175; DB 23; Length 30;

Best Local Similarity 96.7%; Pred. No. 2.8e-13;

Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

Db 1 GSDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 12

AAE20019

ID AAE20019 standard; peptide; 30 AA.

XX AC AAE20019;

XX DT 18-JUN-2002 (first entry)

Complement C1s catalytic site-directed moiety, peptide #20.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.

XX Unidentified.

XX PN WO200198365-A2.

XX DT 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
PT sepsis

XX Disclosure; Page 18; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
XX peptide sequence is complement C1s catalytic site-directed moiety.

XX SQ Sequence 30 AA;

Query Match 94.6%; Score 175; DB 23; Length 30;

Best Local Similarity 95.7%; Pred. No. 2.8e-13;

Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 13

AAE20020

ID AAE20020 standard; peptide; 30 AA.

XX AC AAE20020;

XX DT 18-JUN-2002 (first entry)

Complement C1s catalytic site-directed moiety, peptide #21.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
PT sepsis

XX Disclosure; Page 18; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
XX peptide sequence is complement C1s catalytic site-directed moiety.

XX SQ Sequence 30 AA;

Query Match 94.6%; Score 175; DB 23; Length 30;

Best Local Similarity 96.7%; Pred. No. 2.8e-13;

Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:38:10 ; Search time 24.8571 Seconds
(without alignments)
51.065 Million cell updates/sec

Title: US-09-883-727A-53
Perfect score: 185
Sequence: 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTCNT 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
 - 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
 - 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
 - 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
 - 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	47.6	84	3	US-09-230-180-5
2	85	45.9	50	3	US-08-836-686B-3
3	85	45.9	50	3	US-08-836-686B-4
4	85	45.9	50	3	US-08-836-686B-5
5	85	45.9	50	3	US-08-836-686B-6
6	85	45.9	55	2	US-08-369-829A-1
7	85	45.9	61	3	US-09-230-180-2
8	85	45.9	84	3	US-09-230-180-7
9	85	45.9	84	3	US-09-230-180-8
10	84.5	45.7	119	1	US-08-468-347-20
11	84.5	45.7	119	1	US-08-226-264-24
12	84.5	45.7	119	2	US-08-467-389-20
13	84.5	45.7	119	2	US-08-779-379-20
14	84.5	45.7	119	2	US-08-469-219-20
15	84.5	45.7	119	3	US-09-228-152-19
16	84.5	45.7	136	2	US-08-560-098A-59
17	84	45.4	57	3	US-09-130-121-1
18	84	45.4	57	3	US-09-068-624-1
19	83.5	45.1	55	2	US-08-369-829A-17
20	75.5	40.8	34	2	US-08-369-829A-19
21	75.5	40.8	58	6	5189019-8
22	75.5	40.8	75	6	5189019-4
23	75.5	40.8	136	6	5189019-6
24	74.5	40.3	120	6	5447911-1
25	72	38.9	51	3	US-08-836-686B-2
26	68.5	37.0	94	1	US-08-226-264-22
27	68.5	37.0	118	1	US-08-226-264-23

28	67	36.2	155	1	US-08-468-347-19	Sequence 19, Appl
29	67	36.2	155	2	US-08-467-389-19	Sequence 19, Appl
30	67	36.2	155	2	US-08-779-379-19	Sequence 19, Appl
31	67	36.2	155	2	US-08-469-219-19	Sequence 18, Appl
32	67	36.2	155	3	US-09-228-152-18	Sequence 18, Appl
33	67	36.2	197	1	US-08-468-347-24	Sequence 24, Appl
34	67	36.2	197	2	US-08-467-389-24	Sequence 24, Appl
35	67	36.2	197	2	US-08-779-379-24	Sequence 24, Appl
36	67	36.2	197	2	US-08-469-219-24	Sequence 24, Appl
37	67	36.2	197	3	US-09-228-152-24	Sequence 18, Appl
38	60	32.4	85	1	US-08-468-347-18	Sequence 18, Appl
39	60	32.4	85	2	US-08-467-389-18	Sequence 18, Appl
40	60	32.4	85	2	US-08-779-379-18	Sequence 18, Appl
41	60	32.4	85	2	US-08-469-219-18	Sequence 18, Appl
42	60	32.4	85	2	US-09-228-152-17	Sequence 17, Appl
43	60	32.4	148	1	US-08-468-347-22	Sequence 22, Appl
44	60	32.4	148	2	US-08-467-389-22	Sequence 22, Appl
45	60	32.4	148	2	US-08-779-379-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-230-180-5
; Sequence 5, Application US/69230180
; Patent No. 6183992
; GENERAL INFORMATION:
; APPLICANT: Kim, Sun-Chang
; APPLICANT: Lee, Jae Hyun
; APPLICANT: Kang, Min Hyung
; APPLICANT: Kim, Jeong Hyun
; APPLICANT: Hong, Seung-Suh
; APPLICANT: Lee, Hyun-Soo
; APPLICANT: Samyang Genex Corporation
; APPLICANT: Korea Advanced Institute of Science and Technology
; TITLE OF INVENTION: METHOD FOR MASS PRODUCTION OF
; FILE REFERENCE: 6181/0F135
; CURRENT APPLICATION NUMBER: US/09/230,180
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: PCT/KR98/00132
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: KR 13372/1998
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: KR 21312/1997
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Guamerin/BuforinII fusion protein
US-09-230-180-5

Query Match 47.6%; Score 88; DB 3; Length 84;
Best Local Similarity 51.9%; Pred. No. 0.00029;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
QY 2 CDGFKRLGCTGYGFKTDKKGCEAFCTC 28
| : : : | | | | | | | | | |
DB 32 CTAIRCMTFCPNFGFKVCKNGCEYPTC 58

RESULT 2
US-08-836-686B-3
; Sequence 3, Application US/03836686B
; Patent No. 6239106
; GENERAL INFORMATION:
; APPLICANT: Voerman, Gerard
; TITLE OF INVENTION: A novel family of protease inhibitors, and other

; TITLE OF INVENTION: biologic active substances
; FILE REFERENCE: 70140
; CURRENT APPLICATION NUMBER: US/08/836.686B
; CURRENT FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: EPO 94117053.2
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EPO 95103637.5
; PRIOR FILING DATE: 1995-03-14
; PRIOR APPLICATION NUMBER: PCT/EP95/04223
; PRIOR FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Limnatis nilotica
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(50)
; OTHER INFORMATION: part of a proteinaceous or polypeptide-like
; OTHER INFORMATION: substance from Limnatis nilotica
US-08-836-686B-3

Query Match 45.9%; Score 85; DB 3; Length 50;
Best Local Similarity 48.1%; Pred. No. 0.00042;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTC 28
Db 22 CTPIRCRIYCPKGFVDENGCELPCTC 48

RESULT 3

US-08-836-686B-4
; Sequence 4, Application US/08836686B
; Patent No. 6239106
; GENERAL INFORMATION:
; APPLICANT: Voerman, Gerard
; TITLE OF INVENTION: A novel family of protease inhibitors, and other
; FILE REFERENCE: 70140
; CURRENT APPLICATION NUMBER: US/08/836.686B
; CURRENT FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: EPO 94117053.2
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EPO 95103637.5
; PRIOR FILING DATE: 1995-03-14
; PRIOR APPLICATION NUMBER: PCT/EP95/04223
; PRIOR FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Limnatis nilotica
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(50)
; OTHER INFORMATION: part of a proteinaceous or peptide-like substance
; OTHER INFORMATION: from Limnatis nilotica
US-08-836-686B-4

Query Match 45.9%; Score 85; DB 3; Length 50;
Best Local Similarity 48.1%; Pred. No. 0.00042;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTC 28
Db 22 CTPIRCRIYCPKGFVDENGCELPCTC 48

RESULT 4

US-08-836-686B-5

; Sequence 5, Application US/08836686B
; Patent No. 6239106
; GENERAL INFORMATION:
; APPLICANT: Voerman, Gerard
; TITLE OF INVENTION: A novel family of protease inhibitors, and other
; FILE REFERENCE: 70140
; CURRENT APPLICATION NUMBER: US/08/836.686B
; CURRENT FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: EPO 94117053.2
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EPO 95103637.5
; PRIOR FILING DATE: 1995-03-14
; PRIOR APPLICATION NUMBER: PCT/EP95/04223
; PRIOR FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Limnatis nilotica
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(50)
; OTHER INFORMATION: part of a proteinaceous or peptide-like substance
; OTHER INFORMATION: from Limnatis nilotica
US-08-836-686B-5

Query Match 45.9%; Score 85; DB 3; Length 50;
Best Local Similarity 48.1%; Pred. No. 0.00042;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTC 28
Db 22 CTPIRCRIYCPKGFVDENGCELPCTC 48

RESULT 5

US-08-836-686B-6
; Sequence 6, Application US/08836686B
; Patent No. 6239106
; GENERAL INFORMATION:
; APPLICANT: Voerman, Gerard
; TITLE OF INVENTION: A novel family of protease inhibitors, and other
; FILE REFERENCE: 70140
; CURRENT APPLICATION NUMBER: US/08/836.686B
; CURRENT FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: EPO 94117053.2
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EPO 95103637.5
; PRIOR FILING DATE: 1995-03-14
; PRIOR APPLICATION NUMBER: PCT/EP95/04223
; PRIOR FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Limnatis nilotica
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(50)
; OTHER INFORMATION: part of a proteinaceous or peptide-like substance
; OTHER INFORMATION: from Limnatis nilotica
US-08-836-686B-6

Query Match 45.9%; Score 85; DB 3; Length 50;
Best Local Similarity 48.1%; Pred. No. 0.00042;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTC 28
Db 22 CTPIRCRIYCPKGFVDENGCELPCTC 48

Search completed: August 5, 2003, 09:42:02
Job time : 24.8571 secs

Query Match 45.78; Score 84.5; DB 2; Length 119;

2 CDGFKCRLGLCTYGFKTDKKGCEAFCTC 28
| | : | | : | | : | | : | | :
28 CSGVRCRMHCPHGFORSRYGCE-FCKC 53

RESULT 15
US-09-228-152-19
; Sequence 19, Application US/09228152
; Patent No. 6211341

```

: GENERAL INFORMATION:
: APPLICANT: Zeelon, Elisha P.
: APPLICANT: Werber, Moshe M.
: APPLICANT: Levanon, Avigdor
: TITLE OF INVENTION: NOVEL POLYPEPTIDE HAV
: FILE REFERENCE: 4302aya
: CURRENT APPLICATION NUMBER: US/09/228,152
: CURRENT FILING DATE: 1999-01-11
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patent Ver. 2.0
:

```

EQ ID NO 19
 LENGTH: 119
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:41:00 ; Search time 55.7143 Seconds
(without alignments)
63.948 Million cell updates/sec

Title: US-09-883-727A-53

Perfect score: 185

Sequence: 1 GCDGFKRLGCTGYGFTDKKGCEAFCTCNT 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	30	10 US-09-883-727A-53	Sequence 53, Appl
2	185	100.0	122	10 US-09-883-727A-1	Sequence 1, Appl
3	179	96.8	29	10 US-09-883-727A-121	Sequence 121, Appl
4	175	94.6	30	10 US-09-883-727A-57	Sequence 57, Appl
5	175	94.6	30	10 US-09-883-727A-58	Sequence 58, Appl
6	175	94.6	30	10 US-09-883-727A-59	Sequence 59, Appl
7	175	94.6	30	10 US-09-883-727A-60	Sequence 60, Appl
8	175	94.6	30	10 US-09-883-727A-71	Sequence 71, Appl
9	175	94.6	30	10 US-09-883-727A-72	Sequence 72, Appl
10	175	94.6	30	10 US-09-883-727A-73	Sequence 73, Appl
11	175	94.6	30	10 US-09-883-727A-74	Sequence 74, Appl
12	175	94.6	30	10 US-09-883-727A-85	Sequence 85, Appl
13	175	94.6	30	10 US-09-883-727A-86	Sequence 86, Appl
14	175	94.6	30	10 US-09-883-727A-87	Sequence 87, Appl
15	175	94.6	30	10 US-09-883-727A-88	Sequence 88, Appl

16	165	89.2	30	10 US-09-883-727A-61	Sequence 61, Appl
17	165	89.2	30	10 US-09-883-727A-62	Sequence 62, Appl
18	165	89.2	30	10 US-09-883-727A-63	Sequence 63, Appl
19	165	89.2	30	10 US-09-883-727A-64	Sequence 64, Appl
20	165	89.2	30	10 US-09-883-727A-65	Sequence 65, Appl
21	165	89.2	30	10 US-09-883-727A-66	Sequence 66, Appl
22	165	89.2	30	10 US-09-883-727A-67	Sequence 67, Appl
23	165	89.2	30	10 US-09-883-727A-68	Sequence 68, Appl
24	165	89.2	30	10 US-09-883-727A-69	Sequence 69, Appl
25	165	89.2	30	10 US-09-883-727A-70	Sequence 70, Appl
26	165	89.2	30	10 US-09-883-727A-71	Sequence 71, Appl
27	165	89.2	30	10 US-09-883-727A-72	Sequence 72, Appl
28	165	89.2	30	10 US-09-883-727A-73	Sequence 73, Appl
29	165	89.2	30	10 US-09-883-727A-74	Sequence 74, Appl
30	165	89.2	30	10 US-09-883-727A-75	Sequence 75, Appl
31	165	89.2	30	10 US-09-883-727A-76	Sequence 76, Appl
32	165	89.2	30	10 US-09-883-727A-77	Sequence 77, Appl
33	165	89.2	30	10 US-09-883-727A-78	Sequence 78, Appl
34	165	89.2	30	10 US-09-883-727A-79	Sequence 79, Appl
35	165	89.2	30	10 US-09-883-727A-80	Sequence 80, Appl
36	165	89.2	30	10 US-09-883-727A-81	Sequence 81, Appl
37	165	89.2	30	10 US-09-883-727A-82	Sequence 82, Appl
38	165	89.2	30	10 US-09-883-727A-83	Sequence 83, Appl
39	165	89.2	30	10 US-09-883-727A-84	Sequence 84, Appl
40	165	89.2	30	10 US-09-883-727A-85	Sequence 85, Appl
41	165	89.2	30	10 US-09-883-727A-86	Sequence 86, Appl
42	165	89.2	30	10 US-09-883-727A-87	Sequence 87, Appl
43	165	89.2	30	10 US-09-883-727A-88	Sequence 88, Appl
44	165	89.2	30	10 US-09-883-727A-89	Sequence 89, Appl
45	165	89.2	30	10 US-09-883-727A-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-09-883-727A-53
; Sequence 53, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-53

Query Match 100.0%; Score 185; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 3e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTGYGFTDKKGCEAFCTCNT 30
|||||
Db 1 GCDGFKRLGCTGYGFTDKKGCEAFCTCNT 30

RESULT 2

US-09-883-727A-1
; Sequence 1, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Haementaria ghilianii
US-09-883-727A-1

Query Match 100.0%; Score 185; DB 10; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
||||| ||||||| ||||||| ||||||| |||||||
Db 58 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 87

RESULT 3

US-09-883-727A-121
; Sequence 121, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-121

Query Match 96.8%; Score 179; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
||||| ||||||| ||||||| ||||||| |||||||
Db 1 CDGFKRLGCTYGFKTDKKGCEAFCTCNT 29

RESULT 4

US-09-883-727A-57
; Sequence 57, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 30

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-57

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
||||| ||||||| ||||||| ||||||| |||||||
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 5

US-09-883-727A-58
; Sequence 58, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-58

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
||||| ||||||| ||||||| ||||||| |||||||
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 6

US-09-883-727A-59
; Sequence 59, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-59

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 7

US-09-883-727A-60
; Sequence 60, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-60

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 8

US-09-883-727A-71
; Sequence 71, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-71

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GSDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 9

US-09-883-727A-72
; Sequence 72, Application US/09883727A

; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-72

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 10

US-09-883-727A-73
; Sequence 73, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s catalytic site-directed moiety
US-09-883-727A-73

Query Match 94.6%; Score 175; DB 10; Length 30;
Best Local Similarity 96.7%; Pred. No. 5.8e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

RESULT 11

US-09-883-727A-74
; Sequence 74, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A

Search completed: August 5, 2003, 09:52:46
Job time : 56.2143 secs

THIS PAGE BLANK (USPTO)

Result No.	Query	Score	Match	Length	DB	ID	Description
1	84.5	45.7	120	2	B28806	antistatin B - Mex	
2	84.5	45.7	120	2	A28806	antistatin A - Mex	
3	84	45.4	57	2	A57537	guamerin - Korean	
4	83.5	45.1	119	2	A34398	antistatin - Mexic	
5	83.5	45.1	119	2	S13904	antistatin - Mexic	
6	76	41.1	2150	2	T32497	hypothetical prote	
7	75.5	40.8	136	2	J50209	antistatin precurs	
8	75.5	40.8	220	2	S29195	antistatin - Hydra	
9	74.5	40.3	119	2	A34816	anticoagulant prot	
10	65	35.1	725	2	E96596	hypothetical prote	
11	61.5	33.2	884	2	T18649	hypothetical prote	
12	60	32.4	661	2	B96596	hypothetical prote	
13	57.5	31.1	71	2	A47575	metallothionein 20	
14	57.5	31.1	251	2	A5035	cysteine-rich prot	
15	55.5	30.0	558	2	T17324	hypothetical prote	
16	55	29.7	320	2	T92288	late embryonic abu	
17	55	29.7	707	2	JC2218	procollagen C-endo	
18	55	29.7	1360	2	E96596	hypothetical prote	
19	55	29.7	1680	2	A43434	furin (EC 3.4.21.7	
20	54.5	23.5	71	2	S39420	metallothionein 20	
21	54	23.2	75	2	S17156	metallothionein -	
22	54	23.2	340	2	T3313	hypothetical protei	
23	54	23.2	1464	2	S58984	development protei	
24	53.5	28.9	72	2	S39416	metallothionein 10	
25	53.5	28.9	72	2	S39418	metallothionein 10	
26	53.5	28.9	72	2	S39419	metallothionein 10	
27	53.5	28.9	72	2	S39417	metallothionein 10	
28	53.5	28.9	478	1	SGHUIV	vitronectin precur	
29	53.5	28.9	621	2	I38465	low density lipopr	

C:Superfamily: ultra-high-sulfur keratin

Query Match 40.8%; Score 75.5; DB 2; Length 220;

Best Local Similarity 40.7%; Pred. No. 0.042;

Matches 11; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 CDGFKRLGCTGYGFKTKKGCFAFCYC 28

Db 21 CNKIQCMEFKCFGFGQDENGCD-ICKC 46

RESULT 9

A34816

N:Anticoagulant protein ghilanten - Amazon leech

C:Species: Haementeria ghilianii (Amazon leech)

C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 15-Oct-1994

C:Accession: A34816; A60500; A60953

R:Blankenship, D.T.; Brankamp, R.G.; Manley, G.D.; Cardin, A.D.

Biochem. Biophys. Res. Commun. 166, 1384-1389, 1990

A:Title: Amino acid sequence of ghilanten: anticoagulant-antimetastatic principle of the

A:Reference number: A34816; MUID:90165947; PMID:2306252

A:Accession: A34816

A:Molecule type: protein

A:Residues: 1-119 <BLA>

R:Condra, C.; Nutt, E.; Petroski, C.J.; Simpson, E.; Friedman, P.A.; Jacobs, J.W.

Thromb. Haemost. 61, 437-441, 1989

A:Title: Isolation and structural characterization of a potent inhibitor of coagulation

A:Reference number: A60500; MUID:90020158; PMID:2572073

A:Accession: A60500

A:Molecule type: protein

A:Residues: 1-9 <BRA>

R:Brankamp, R.G.; Blankenship, D.T.; Sunkara, P.S.; Cardin, A.D.

J. Lab. Clin. Med. 115, 89-97, 1990

A:Title: Ghilantens: anticoagulant-antimetastatic proteins from the South American leech

A:Reference number: A60953; MUID:90132188; PMID:2299260

A:Accession: A60953

A:Molecule type: protein

A:Residues: 1-9 <BRA>

C:Superfamily: antistasin

C:Keywords: anticoagulant; heparin binding; pyroglutamic acid

F:/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:/Inhibitory site: Arg (coagulation factor Xa, trypsin) #status experimental

Query Match 40.3%; Score 74.5; DB 2; Length 119;

Best Local Similarity 41.4%; Pred. No. 0.035;

Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 2 CDGFKRLGCTGYGFKTKKGCFAFCYC 30

Db 28 CPVRCRVYCSHGFRSRYGCV-CHCRT 55

RESULT 10

E96596

hypothetical protein T5A14.16 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: E96596

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96596

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-725 <STO>

A:Cross-references: GB:AE005173; NID:g4204272; PIDN:AA010653.1; GSPDB:GN00141

C:Genetics:

A:Gene: T5A14.16

A:Map position: 1

Query Match 35.1%; Score 65; DB 2; Length 725;

Best Local Similarity 44.4%; Pred. No. 1.9;

Matches 12; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 3 DGFKC----RLGCTGYGFKTKKGCFAF 25

Db 468 DIFKCGICGIGCGFYKCDKNCDEF 494

RESULT 11

T18649

hypothetical protein B0024.14 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18649

R:McMurray, A.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z19001

A:Accession: T18649

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-884 <WIL>

A:Cross-references: EMBL:Z11178; PIDN:CAA94886.1; GSPDB:GN00023; CESP:B0024.14

A:Experimental source: clone B0024

C:Genetics:

A:Gene: CESP:B0024.14

A:Map position: 5

A:Introns: 46/3; 84/1; 212/1; 307/2; 345/2; 394/1; 424/2; 481/1; 596/1; 702/1; 765/3;

Query Match 33.2%; Score 61.5; DB 2; Length 884;

Best Local Similarity 43.5%; Pred. No. 5.7;

Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 6 KRLGCTGYGFKTKKGCFAFCYC 28

Db 371 OCHKHCLYGFETNSAGC-SLCKC 392

RESULT 12

B96596

hypothetical protein T1813.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96596

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96596

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-661 <STO>

A:Cross-references: GB:AE005173; NID:g11094790; PIDN:AAG29722.1; GSPDB:GN00141

C:Genetics:

A:Gene: T1813.4

A:Map position: 1

Query Match 32.4%; Score 60; DB 2; Length 661;

Best Local Similarity 40.7%; Pred. No. 7;

Matches 11; Conservative 6; Mismatches 6; Indels 4; Gaps 1;

QY 3 DGFKC-----RLGCTYGFKTDKKGCEAF 25
| | | | | : | | | : | | | : |
Db 429 DIFQIGGGRIGCGGFFKCKDEKDCDEF 455

RESULT 13
S47576
metallothionein 20-Ib - blue mussel
C:Species: Mytilus edulis (blue mussel)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-1999
C:Accession: S47576
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.; F.
Eur. J. Biochem. 218, 183-194, 1993
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of metal
A:Reference number: S394116; MUID:94062828; PMID:8243463
A:Accession: S47576
A:Molecule type: protein
A:Residues: 1-71 <MAC>
C:Superfamily: metallothionein
C:Keywords: chelation; metal binding

Query Match 31.1%; Score 57.5; DB 2; Length 71;
Best Local Similarity 38.9%; Pred. No. 2.7;
Matches 14; Conservative 3; Mismatches 10; Indels 9; Gaps 2;

QY 1 GCDGFK-----CRLGCTYGFKTDKKGCEAFCTC 28
| | | | | : | | | : | | | : |
Db 36 GCGCKVVKCGTGCKGCDCTGPTNCK-CEAGCSC 70

RESULT 14
A55035
cysteine-rich protein CRP1 - earthworm (Enchytraeus buchholzi)
C:Species: Enchytraeus buchholzi
C:Date: 14-Nov-1994 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: A55035; S45034
R:Willuhn, J.; Schmitt-Wrede, H.P.; Greven, H.; Wunderlich, F.
J. Biol. Chem. 269, 24688-24691, 1994
A:Title: cDNA cloning of a cadmium-inducible mRNA encoding a novel cysteine-rich, non-me
A:Reference number: A55035; MUID:95014230; PMID:7929141
A:Accession: A55035
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-251 <WIL>
A:Cross-references: EMBL:X79344; NID:g488802; PIDN:CAA55899.1; PID:g488803
C:Superfamily: ultra-high-sulfur keratin

Query Match 31.1%; Score 57.5; DB 2; Length 251;
Best Local Similarity 41.4%; Pred. No. 6.8;
Matches 12; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTCNT 30
| : | | | | | | | | | | | : |
Db 50 CEKGECKRGCKGCGCTPCGVEG-CPCGS 77

RESULT 15
T17324
hypothetical protein DKFZp564P2063.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17324
R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18727
A:Accession: T17324
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-558 <DUE>
A:Cross-references: EMBL:AL117610
A:Experimental source: fetal brain; clone DKFZp564P2063
C:Genetics:

A:Note: DKFZp564P2063.1
Query Match 30.0%; Score 55.5; DB 2; Length 558;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
QY 8 RLGCTYGFKTDKKG-CEAFC 26
: | | | | : | | | |
Db 51 KLACCYGWRNRNSKGVCEATC 70
Search completed: August 5, 2003, 09:41:27
Job time : 23.2857 secs

AC P82354;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Theromycin (Thrombin inhibitor).
OS Theromycin tessulatum (Leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Rhynchobdellida; Glossiphoniidae; Theronomyon.
OX NCBI_TaxID=13286;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Head;
RX MEDLINE=20469422; PubMed=10837466;
RA Salzet M., Chopin V., Baert J.-L., Matias I., Malecha J.;
RT "Theromycin, a novel leech thrombin inhibitor.";
RL J. Biol. Chem. 275:30774-30780(2000).
CC -|- FUNCTION: POTENT THROMBIN-SPECIFIC INHIBITOR.
CC -|- SUBUNIT: Homodimer.
CC -|- PTM: EIGHT DISULFIDE BONDS ARE PRESENT.
CC -|- MASS SPECTROMETRY: MW=14491; METHOD=MALDI.
CC -|- SIMILARITY: BELONGS TO THE ANTISTASIN FAMILY.
KW Serine protease inhibitor.
FT ACT_SITE 36 37 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 67 AA; 7257 MW; 1956FE93A57B4FA CRC64;

Query Match . 47.0%; Score 87; DB 1; Length 67;
Best Local Similarity 53.3%; Pred. No. 6.2e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 6; Gaps 2;

QY 1 GCDGFKRLGCTGYGFKTDKKGCEAFCTC 30
III :II I : : III:IIIIII
DB 29 GCDDAQCR--CS----SDANGCESFCTC 52

RESULT 3
ANTA_HIRME STANDARD; PRT; 55 AA.
AC P80302;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hirustasin.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=94155900; PubMed=8112345;
RA Soellner C., Mentele R., Eckerskorn C., Fritz H., Sommerhoff C.P.;
RT "Isolation and characterization of hirustasin, an antistasin-type
serine-proteinase inhibitor from the medicinal leech Hirudo
medicinalis.";
RL Eur. J. Biochem. 219:937-943(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH KALLIKREIN.
RX MEDLINE=97184690; PubMed=9032072;
RA Mittl P.R., di Marco S., Fendrich G., Pohlig G., Heim J.,
Mittl P.R., di Marco S., Fendrich G., Prietle J.P., Grutter M.G.;
RA Sommerhoff C., Fritz H., Priestle J.P., Grutter M.G.;
RT "A new structural class of serine protease inhibitors revealed by the
structure of the hirustasin-kallikrein complex.";
RL Structure 5:253-264(1997).
RN [3]
RP ERRATUM.
RA Mittl P.R., di Marco S., Fendrich G., Pohlig G., Heim J.,
Sommerhoff C., Fritz H., Priestle J.P., Grutter M.G.;
RL Structure 5:585-585(1997).
CC -|- FUNCTION: ACTS AS AN INHIBITOR OF TISSUE KALLIKREIN, TRYPSIN,
CC CHYMOTRYPSIN AND NEUTROPHIL CATHEPSIN G.
CC -|- SIMILARITY: BELONGS TO THE ANTISTASIN FAMILY.
DR PDB: 1HIA; 24-DEC-97.
DR PDB: 1BX7; 27-APR-99.
DR PDB: 1BX8; 27-APR-99.

DR InterPro; IPR004094; Antistasin.
DR Pfam; PF02822; Antistasin; 1.
KW Serine protease inhibitor; Heparin-binding; 3D-structure.
FT ACT_SITE 30 31 REACTIVE BOND.
FT DISULFID 6 17
FT DISULFID 11 22
FT DISULFID 24 44
FT DISULFID 29 48
FT DISULFID 33 50
FT TURN 6 7
FT TURN 13 14
FT STRAND 15 18
FT TURN 19 20
FT STRAND 21 24
FT STRAND 27 29
FT STRAND 37 39
FT TURN 41 42
FT STRAND 45 50
SQ SEQUENCE 55 AA; 5878 MW; 234226BEE3CC532E CRC64;

Query Match 45.9%; Score 85; DB 1; Length 55;
Best Local Similarity 48.1%; Pred. No. 9.5e-05;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 CDGFKRLGCTGYGFKTDKKGCEAFCTC 28
I : II : III : III : I :
DB 24 CNEVHCRLRCKYGLKKDENGCEYPCSC 50

RESULT 4
ANTA_HAEOF STANDARD; PRT; 136 AA.
AC P15358; O9TW08; O9TX45;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Antistasin precursor (ATS) (Blood coagulation factor Xa/proclotting
enzyme inhibitor).
OS Haemeteria officinalis (Mexican leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Rhynchobdellida; Glossiphoniidae; Haementeria.
OX NCBI_TaxID=6410;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89252921; PubMed=2470652;
RA Han J.H., Law S.W., Keller P.M., Kniskern P.J., Silberklang M.,
Tung J.S., Gasic T.B., Gasic G.J., Friedman P.A., Ellis R.W.;
RT "Cloning and expression of cDNA encoding antistasin, a leech-derived
protein having anti-coagulant and anti-metastatic properties.";
RL Gene 75:47-57(1989).
RN [2]
RP SEQUENCE OF 18-136.
RC TISSUE=Saliva;
RX MEDLINE=86273105; PubMed=3164720;
RA Nutt E., Gasic T., Rodkey J., Gasic G.J., Jacobs J.W., Friedman P.A.,
Simpson E.;
RT "The amino acid sequence of antistasin. A potent inhibitor of factor
Xa reveals a repeated internal structure.";
RL J. Biol. Chem. 263:10162-10167(1988).
RN [3]
RP SEQUENCE OF 18-136.
RC TISSUE=Saliva;
RX MEDLINE=9409722; PubMed=8271959;
RA Dunwiddie C.T., Waxman L., Vlasuk G.P., Friedman P.A.;
RT "Purification and characterization of inhibitors of blood coagulation
factor Xa from hemaphysal organisms.";
RL Meth. Enzymol. 223:291-312(1993).
RN [4]
RP REACTIVE SITE.
RX MEDLINE=89380295; PubMed=2777803;
RA Dunwiddie C., Thornberry N.A., Bull H.G., Sardana M., Friedman P.A.,
Jacobs J.W., Simpson E.;
RT "Antistasin, a leech-derived inhibitor of factor Xa. Kinetic analysis


```

SQ SEQUENCE 57 AA; 6123 MW; C5116B6E5481D7B CRC64;
Query Match 45.48; Score 84; DB 1; Length 57;
Best Local Similarity 48.18; Pred. No. 0.00013;
Matches 13; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Oy 2 CDGFKRLGCTGYGFKTDKKGCEAFCTC 28
   : : : : : : : : : : : : : : : :
Db 30 CTAIRCMIFCNGFKVDNGCEYPTC 56

RESULT 6
ANTA_HYDMA STANDARD; PRT; 220 AA.
AC P38977;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antistatin precursor (ATS) (Blood coagulation factor Xa/proclotting
DE enzyme inhibitor).
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF1;
RX MEDLINE=92387373; PubMed=1516699;
RA Holstein T.W., Mala C., Kurz E., Bauer K., Greber M., David C.N.;
RT "The primitive metazoan Hydra expresses antistatin, a serine protease
RT inhibitor of vertebrate blood coagulation: cDNA cloning, cellular
RT localisation and developmental regulation.";
RL FEBS Lett. 309:288-292(1992).
CC -1- FUNCTION: THIS HIGHLY DISULFIDE-BONDED PROTEIN IS A POTENT
CC INHIBITOR OF FACTOR XA. FACILITATES DIGESTION OF TISSUES
CC AND MAY ALSO PROTECT THE GASTRIC TISSUES FROM ITS OWN DIGESTIVE
CC ENZYMES. MAY HAVE THERAPEUTIC UTILITY AS AN ANTICOAGULANT. ALSO
CC EXHIBITS A STRONG METASTATIC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: GLAND CELLS. IT IS MORE STRONGLY EXPRESSED
CC IN THE HEAD THAN IN THE GASTRIC TISSUE.
CC -1- SIMILARITY: BELONGS TO THE ANTISTATIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X67590; CAA47864.1; -.
CC PIR; S29195; S29195.
CC HSP; P15358; 1SKZ.
CC InterPro: IPR004094; Antistatin.
CC Pfam; PF02822; Antistatin; 6.
CC Serine protease inhibitor; Repeat; Heparin-binding;
CC Blood coagulation; Signal.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 220 ANTISTATIN.
CC DOMAIN 21 220 6 X APPROXIMATE TANDEM REPEATS.
CC REPEAT 21 53 1.
CC REPEAT 54 90 2.
CC REPEAT 91 119 3.
CC REPEAT 120 153 4.
CC REPEAT 154 182 5.
CC REPEAT 183 220 6.
CC ACT_SITE 27 28 REACTIVE BOND (BY SIMILARITY).
CC ACT_SITE 60 61 REACTIVE BOND (BY SIMILARITY).
CC ACT_SITE 98 99 REACTIVE BOND (BY SIMILARITY).
CC ACT_SITE 161 162 REACTIVE BOND (BY SIMILARITY).
CC SEQUENCE 220 AA; 25016 MW; E7987F01900D0278 CRC64;

```

```

Query Match 40.88; Score 75.5; DB 1; Length 220;
Best Local Similarity 40.78; Pred. No. 0.0053;
Matches 11; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

Oy 2 CDGFKRLGCTGYGFKTDKKGCEAFCTC 28
   : : : : : : : : : : : : : : : :
Db 21 CNKIQCRMFCKFGFQDNGCD-ICK 46

RESULT 7
ANTA_HAEGH STANDARD; PRT; 119 AA.
AC P16242;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ghilanten.
OS Haementeria ghilianii (Amazon leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Rhyachobdellida; Glossiphoniidae; Haementeria.
OX NCBI_TaxID=6409;
RN [1]
RP SEQUENCE.
RC TISSUE=Saliva;
RX MEDLINE=90165947; PubMed=2306252;
RA Blankenship D.T., Brankamp R.G., Manley G.D., Cardin A.D.;
RT "Amino acid sequence of ghilanten: anticoagulant-antimetastatic
RT principle of the South American leech, Haementeria ghilianii.";
RL Biochem. Biophys. Res. Commun. 166:1384-1389(1990).
CC -1- FUNCTION: THIS HIGHLY DISULFIDE-BONDED PROTEIN IS A POTENT
CC INHIBITOR OF FACTOR XA. MAY HAVE THERAPEUTIC UTILITY AS AN
CC ANTICOAGULANT. ALSO EXHIBITS A STRONG METASTATIC ACTIVITY.
CC -1- MISCELLANEOUS: BINDS TO HEPARIN-AGAROSE, BINDS TO SULFATED
CC GLYCOCONGUGATES.
CC -1- SIMILARITY: BELONGS TO THE ANTISTATIN FAMILY.
CC PIR; A34816; A34816.
CC HSP; P15358; 1SKZ.
CC InterPro: IPR004094; Antistatin.
CC Pfam; PF02822; Antistatin; 2.
CC Serine protease inhibitor; Repeat; Heparin-binding;
CC Pyroliadone carboxylic acid.
CC MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
CC DOMAIN 2 110 2 X APPROXIMATE TANDEM REPEATS.
CC REPEAT 2 55 1.
CC REPEAT 56 110 2.
CC DOMAIN 97 100 HEPARIN-BINDING (POTENTIAL).
CC DOMAIN 111 118 HEPARIN-BINDING (POTENTIAL).
CC ACT_SITE 34 35 REACTIVE BOND (BY SIMILARITY).
CC ACT_SITE 89 90 REACTIVE BOND (BY SIMILARITY).
CC DISULFID 8 19 BY SIMILARITY.
CC DISULFID 13 26 BY SIMILARITY.
CC DISULFID 28 48 BY SIMILARITY.
CC DISULFID 33 51 BY SIMILARITY.
CC DISULFID 37 53 BY SIMILARITY.
CC DISULFID 62 73 BY SIMILARITY.
CC DISULFID 67 80 BY SIMILARITY.
CC DISULFID 82 103 BY SIMILARITY.
CC DISULFID 88 106 BY SIMILARITY.
CC DISULFID 92 108 BY SIMILARITY.
CC SEQUENCE 119 AA; 13317 MW; 5A94805DBB850EF CRC64;
Query Match 40.38; Score 74.5; DB 1; Length 119;
Best Local Similarity 41.48; Pred. No. 0.0042;
Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

Oy 2 CDGFKRLGCTGYGFKTDKKGCEAFCTC 30
   : : : : : : : : : : : : : : : :
Db 28 CPEVRCRVYCSHGFSRSGEV-CRCRT 55

RESULT 8
PIGU_HIRNI STANDARD; PRT; 48 AA.
ID PIGU_HIRNI
```

AC P81499;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pigamerin.
 OS Hirudo nipponia (Leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 ON NCBI_TaxID=42736;
 RX [1]
 RP SEQUENCE.
 RX MEDLINE=98351570; PubMed=9688284;
 RA Kim D.R., Kang K.W.;
 RT "Amino acid sequence of pigamerin, an antistatin-type protease
 RT inhibitor from the blood sucking leech Hirudo nipponia.";
 RL Eur. J. Biochem. 254:692-697(1998).
 CC -1- FUNCTION: INHIBITS PLASMA AND TISSUE KALLIKREIN, AND TRYPSIN. MAY
 CC BE INVOLVED IN LEECH HEMATOPHAGIA.
 CC -1- SIMILARITY: BELONGS TO THE ANTISTATIN FAMILY.
 CC HSSP; P80302; IHIA.
 DR InterPro: IPR004094; Antistatin.
 DR Pfam: PF02822; Antistatin; 1.
 KW Serine protease inhibitor.
 FT ACT_SITE 27 28 REACTIVE BOND.
 FT DISULFID 3 14 BY SIMILARITY.
 FT DISULFID 8 19 BY SIMILARITY.
 FT DISULFID 21 41 BY SIMILARITY.
 FT DISULFID 26 45 BY SIMILARITY.
 FT DISULFID 30 47 BY SIMILARITY.
 SQ SEQUENCE 48 AA; 5101 MW; 0331C2D807F32FDC CRC64;
 Query Match 37.8%; Score 70; DB 1; Length 48;
 Best Local Similarity 52.2%; Pred. No. 0.0074;
 Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 OY 6 KRLGCTYGFKTKGCEAFRC 28
 DB 25 QCRKYPNGFKKDKGCTFPCTC 47
 RESULT 9
 TLD_BRARE STANDARD; PRT; 1022 AA.
 AC 057460;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dorsal-ventral patterning tolloid protein precursor (EC 3.4.24.-)
 DE (Mini fin protein).
 GN TOLLOID OR TLD OR MFN.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 ON NCBI_TaxID=7955;
 RX [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=gastrula;
 RX MEDLINE=98057457; PubMed=9395394;
 RA Blader P., Rastegar S., Fischer N., Straehle U.;
 RT "Cleavage of the BMP-4 antagonist chordin by zebrafish Tolloid.";
 RL Science 278:1937-1940(1997).
 RN [2]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=99307076; PubMed=10375503;
 RA Connors S.A., Trout J., Ekker M., Mullins M.C.;
 RT "The role of tolloid/mini fin in dorsoventral pattern formation of the
 RT zebrafish embryo.";
 RL Development 126:3119-3130(1999).
 CC -1- FUNCTION: Required for patterning ventral tissues of the tail. May
 CC increase bone morphogenetic protein (BMP) activity at the end of
 CC gastrulation by proteolytic cleavage of chordin and release of BMP
 CC from inactive complexes.

CC -1- TISSUE SPECIFICITY: During gastrulation, accumulates around the
 CC closing blastopore with greater expression ventrally. At the
 CC animal pole, expressed in the ectoderm flanking the anterior
 CC neural plate. At the 10-somite stage, expressed in the developing
 CC tailbud and cranial neural crest. At the 20-somite stage, also
 CC expressed in the hematopoietic system.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 5 CUB domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF027596; AAC60304.1;
 CC HSSP; P35555; LEIN.
 DR MEROPS; M12.016;
 DR ZFIN; ZDB-GENE-990415-265; tolloid.
 DR InterPro: IPR001506; Astacin.
 DR InterPro: IPR000152; ASX_hydroxyl.
 DR InterPro: IPR000859; CUB_domain.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR006209; EGF_like.
 DR InterPro: IPR006026; Zn_Mtpeptidse.
 DR InterPro: IPR006025; Zn_Mtpeptidse.
 DR Pfam; PF01400; Astacin; 1.
 DR Pfam; PF00431; CUB; 5.
 DR Pfam; PF00008; EGF; 2.
 DR PRINTS; PR00480; ASTACIN.
 DR SMART; SM00042; CUB; 5.
 DR SMART; SM00179; EGF_CA; 2.
 DR SMART; SM00235; ZnMC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 DR PROSITE; PS01180; CUB; 5.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Developmental protein; Hydrolase; Protease; Metalloprotease; Zinc;
 KW Metal-binding; Calcium; EGF-like domain; Repeat; Signal; Glycoprotein;
 KW Zymogen.
 FT SIGNAL 1 32 POTENTIAL.
 FT PROPEP 33 156 POTENTIAL.
 FT CHAIN 157 1022 DORSAL-VENTRAL PATTERNING TOLLOID
 FT PROTEIN.
 FT DOMAIN 157 357 METALLOPROTEASE (BY SIMILARITY).
 FT DOMAIN 358 470 CUB 1.
 FT DOMAIN 471 583 CUB 2.
 FT DOMAIN 583 624 EGF-LIKE 1, CALCIUM-BINDING
 FT (POTENTIAL).
 FT DOMAIN 627 739 CUB 3.
 FT DOMAIN 739 779 EGF-LIKE 2, CALCIUM-BINDING
 FT (POTENTIAL).
 FT DOMAIN 783 895 CUB 4.
 FT DOMAIN 896 1012 CUB 5.
 FT METAL 249 249 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 250 250 BY SIMILARITY.
 FT METAL 253 253 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 259 259 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 358 384 BY SIMILARITY.
 FT DISULFID 411 433 BY SIMILARITY.
 FT DISULFID 471 497 BY SIMILARITY.
 FT DISULFID 524 546 BY SIMILARITY.
 FT DISULFID 567 599 BY SIMILARITY.
 FT DISULFID 595 608 BY SIMILARITY.
 FT DISULFID 610 623 BY SIMILARITY.
 FT DISULFID 627 653 BY SIMILARITY.
 FT DISULFID 680 702 BY SIMILARITY.
 FT DISULFID 743 754 BY SIMILARITY.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN RASK L., MORSE B.;
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96305376; PubMed=8706697;
RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,
RA Rask L., Morse B.;
RT Cloning and sequencing of human gp330, a Ca(2+)-binding receptor
RT with potential intracellular signaling properties.";
RL Eur. J. Biochem. 239:132-137(1996).
RN [2]
RN SEQUENCE OF 2705-4453 FROM N.A.
RC TISSUE=Kidney;
RA Knaak C., Argaves W.S.;
RN Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 3833-4453 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95048397; PubMed=7959795;
RA Korenberg J.R., Argaves W.S.;
RA Strickland D.K., Argaves W.S.;
RT "Chromosomal localization of human genes for the LDL receptor family
RT member glycoprotein 330 (LRP2) and its associated protein RAP
RT (LRPAP1).";
RL Genomics 22:88-93(1994).
RN [4]
RN SEQUENCE OF 4139-4406 FROM N.A.
RX MEDLINE=94244704; PubMed=8187828;
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,
RA Klareskog L., Akerstrom G., Rask L.;
RT "A protein involved in calcium sensing of the human parathyroid and
RT placental cytotrophoblast cells belongs to the LDL-receptor protein
RT superfamily.";
RL Exp. Cell Res. 212:344-350(1994).
RN [5]
RN FUNCTION.
RX MEDLINE=95286588; PubMed=7768901;
RA Kounas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,
RA Brewer B.H., Strickland D.K., Argaves W.S.;
RT "Identification of glycoprotein 330 as an endocytic receptor for
RT apolipoprotein J/clusterin.";
RL J. Biol. Chem. 270:13070-13075(1995).
CC -1- FUNCTION: BINDS SPECIFICALLY CLUSTERIN WITH HIGH AFFINITY. BUT
CC ALSO LIGANDS IN COMMON WITH OTHER FAMILY MEMBERS: PLASMINOGEN,
CC EXTRACELLULAR MATRIX COMPONENTS, PLASMINOGEN ACTIVATOR-PLASMINOGEN
CC ACTIVATOR INHIBITOR TYPE 1 COMPLEX, APOLOPROTEIN E-ENRICHED
CC BETA-VLDL, LIPOPROTEIN LIPASE, LACTOFERRIN AND CALCITRIOL.
CC -1- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B (BY SIMILARITY).
CC -1- FUNCTION: MAY PARTICIPATE IN REGULATION OF PARATHYROID-HORMONE AND
CC PARA-THYROID-HORMONE-RELATED PROTEIN RELEASE.
CC -1- SUBUNIT: Forms a multimeric complex together with a receptor-
CC associated protein (RAP). Binds to ankyrin-repeat family A protein
CC 2 (ANKRA2) (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: ABSORPTIVE EPITHELIA, INCLUDING RENAL
CC PROXIMAL TUBULES.
CC -1- SIMILARITY: Contains 36 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 37 LDL-receptor class B domains.
CC -1- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U33837; AAB41649.1; -;
DR EMBL; U04441; AAB02882.1; -;
DR EMBL; S73145; AAB30825.1; -;
DR PIR; I53413; I53413.
DR HSP; Q07954; ICR8.
DR Genew; HGNC:6694; LRP2.
DR MIM; 600073; -;
DR GO; GO:0005764; C:lysosome; TAS.
DR GO; GO:0006629; P:lipid metabolism; TAS.
DR GO; GO:0006486; P:protein amino acid glycosylation; TAS.
DR GO; GO:0006998; P:receptor mediated endocytosis; TAS.
DR InterPro; IPR000152; ASX_HYDROXYL.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF_13.
DR Pfam; PF00037; ldl_recept_a; 36.
DR Pfam; PF00058; ldl_recept_b; 37.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLa; 36.
DR SMART; SM00135; LY; 33.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 31.
DR PROSITE; PS00068; LDLRA_2; 36.
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
KW Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 4655
FT DOMAIN 26 4423
FT TRANSMEM 4424 4446
FT DOMAIN 4447 4655
FT DOMAIN 26 64
FT DOMAIN 65 105
FT DOMAIN 145 181
FT DOMAIN 182 219
FT DOMAIN 220 258
FT DOMAIN 264 308
FT DOMAIN 309 347
FT DOMAIN 348 386
FT DOMAIN 436 477
FT DOMAIN 479 520
FT DOMAIN 522 567
FT DOMAIN 569 612
FT DOMAIN 613 653
FT DOMAIN 659 705
FT DOMAIN 753 794
FT DOMAIN 796 836
FT DOMAIN 838 880
FT DOMAIN 882 924
FT DOMAIN 970 1014
FT DOMAIN 1024 1062
FT DOMAIN 1065 1103
FT DOMAIN 1107 1145
FT DOMAIN 1147 1185
FT DOMAIN 1186 1224
FT DOMAIN 1228 1268
FT DOMAIN 1269 1307
FT DOMAIN 1310 1350
FT DOMAIN 1349 1389
FT DOMAIN 1390 1429
FT DOMAIN 1478 1519
FT DOMAIN 1521 1562
FT DOMAIN 1566 1608
FT DOMAIN 1610 1653
FT DOMAIN 1655 1695
FT DOMAIN 1700 1741
FT DOMAIN 1790 1831
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
LDL-RECEPTOR CLASS A 1.
LDL-RECEPTOR CLASS A 2.
LDL-RECEPTOR CLASS A 3.
LDL-RECEPTOR CLASS A 4.
LDL-RECEPTOR CLASS A 5.
LDL-RECEPTOR CLASS A 6.
LDL-RECEPTOR CLASS A 7.
EGF-LIKE 1.
EGF-LIKE 2.
LDL-RECEPTOR CLASS B 1.
LDL-RECEPTOR CLASS B 2.
LDL-RECEPTOR CLASS B 3.
LDL-RECEPTOR CLASS B 4.
LDL-RECEPTOR CLASS B 5.
EGF-LIKE 3.
LDL-RECEPTOR CLASS B 6.
LDL-RECEPTOR CLASS B 7.
LDL-RECEPTOR CLASS B 8.
LDL-RECEPTOR CLASS B 9.
EGF-LIKE 4.
LDL-RECEPTOR CLASS A 8.
LDL-RECEPTOR CLASS A 9.
LDL-RECEPTOR CLASS A 10.
LDL-RECEPTOR CLASS A 11.
LDL-RECEPTOR CLASS A 12.
LDL-RECEPTOR CLASS A 13.
LDL-RECEPTOR CLASS A 14.
LDL-RECEPTOR CLASS A 15.
EGF-LIKE 5.
EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
LDL-RECEPTOR CLASS B 10.
LDL-RECEPTOR CLASS B 11.
LDL-RECEPTOR CLASS B 12.
LDL-RECEPTOR CLASS B 13.
LDL-RECEPTOR CLASS B 14.
EGF-LIKE 7.
LDL-RECEPTOR CLASS B 15.


```

CC -----
DR EMBL; X59862; CAA4252.1; -
DR PIR; S17156; S17156.
DR HSSP; P55949; 1DME.
DR InterPro; IPR003019; Metallthion.
DR InterPro; IPR001008; Metallthion_2.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00875; MTMOLJUSC.
KW Metal-binding; Metal-thiolate cluster; Cadmium; Acetylation.
FT INT_MET 0 0
FT MOD_RES 1 1 ACETYLATION (PARTIAL).
FT SEQUENCE 74 AA; 7213 MW; 6364DFB4046E205E CRC64;
SQ
Query Match 29.28; Score 54; DB 1; Length 74;
Best Local Similarity 37.08; Pred. No. 1.3;
Matches 10; Conservative 3; Mismatches 14; Indels 0; Gaps 0;
QY 4 GFCRLGCTGYGFKTKKGCCEAFCTCNT 30
Db 22 GCRGPGCKGDDCKCAGCKVKCSCTS 48
RESULT 15
SORL_MOUSE
ID SORL_MOUSE STANDARD; PRT; 2215 AA.
AC O88307; O54711; O70581;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sortilin-related receptor precursor (Sorting protein-related receptor
DE containing LDLR class A repeats) (mSorLA) (SorLA-1) (Low-density
DE lipoprotein receptor relative with 11 ligand-binding repeats) (LDLR
DE relative with 11 ligand-binding repeats) (LR11) (Gp250).
GN SORL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98392848; PubMed=9726247;
RA Kanaki T., Bujo H., Hirayama S., Tanaka K., Yamazaki H., Seimiya K.,
RA Morisaki N., Schneider W.J., Saito Y.;
RA "Developmental regulation of Lrll expression in murine brain.";
RL DNA Cell Biol. 17:647-657(1998).
RN [2]
RP SEQUENCE OF 183-2215 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98168844; PubMed=9510025;
RA Hermans-Borgmeyer I., Hampe W., Schinke B., Methner A., Nykjaer A.,
RA Suesens U., Fenger U., Herbarth B., Schaller H.C.;
RA "Unique expression pattern of a novel mosaic receptor in the
RA developing cerebral cortex.";
RL Mech. Dev. 70:65-76(1998).
RN [3]
RP SEQUENCE OF 1119-1713 FROM N.A.
RC STRAIN=Swiss Webster;
RA Boehmelt G., Antonio L., Iscove N.N.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: LIKELY TO BE A MULTIFUNCTIONAL ENDOCYTIC RECEPTOR, THAT
CC MAY BE IMPLICATED IN THE UPTAKE OF LIPOPROTEINS AND OF PROTEASES.
CC BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN OF PLASMA,
CC AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS. BINDS THE RECEPTOR-
CC ASSOCIATED PROTEIN (RAP). COULD PLAY A ROLE IN CELL-CELL
CC INTERACTION. MAY PLAY A ROLE IN NEURAL ORGANIZATION, AS WELL AS
CC THE ESTABLISHMENT OF EMBRYONIC ORGAN SYSTEMS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, WHERE IT IS MAINLY
CC EXPRESSED IN ADULT CEREBELLUM, HIPPOCAMPAL CA REGIONS, DENTATE
CC GYRUS, AND TO A MUCH LESSER EXTENT IN THE CEREBRAL CORTEX.
CC DETECTABLE IN KIDNEY, SKELETAL MUSCLE, LUNG AND SPLEEN, BUT NOT IN
CC THE LIVER.

```

```

CC -----
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AS EARLY AS EMBRYONIC DAY 6.5
CC (E6.5) AND PEAKS AT E11, THE MAIN LOCATION IS IN THE CNS DURING
CC DEVELOPMENT. AT EARLY STAGES, IT IS ABUNDANT IN A SUBPOPULATION OF
CC NEURONS IN THE CEREBRAL CORTEX, IN THE HIPPOCAMPUS, AND GRANULAR
CC AND PURKINJE CELL LAYERS IN THE CEREBELLUM, WHEREAS IN THE ADULT,
CC EXPRESSION IN CEREBELLAR GRANULAR CELLS AND IN THE CEREBRAL CORTEX
CC IS LOW. EXPRESSION OCCURS ALSO IN A VARIETY OF GLANDS AND ORGANS
CC DURING ORGANOGENESIS.
CC -1- PIM: THE PROPEPTIDE REMOVED IN THE N-TERMINUS MAY BE CLEAVED BY
CC FURIN OR HOMOLOGOUS PROTEASES (BY SIMILARITY).
CC -1- SIMILARITY: Contains 5 BNR repeats.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 11 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB015790; BAA31219.1; -
CC EMBL; AF031816; AAC16739.1; -
CC EMBL; Y12004; CAA72732.1; -
CC PIR; T00348; T00348.
CC HSSP; P01130; 1AJJ.
CC MGD; MGI-1202296; Sorl1.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR002860; GH_BNR.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; LDL_receptor_rep.
CC InterPro; IPR006581; VPS10.
CC Pfam; PF02012; BNR; 5.
CC Pfam; PF00041; fn3; 4.
CC Pfam; PF00057; ldl_recept_a; 11.
CC Pfam; PF00058; ldl_recept_b; 5.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00060; FN3; 4.
CC SMART; SM00192; LDLA; 11.
CC SMART; SM00135; LY; 5.
CC SMART; SM00602; VPS10; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01209; LDLRA_1; 10.
CC PROSITE; PS00068; LDLRA_2; 11.
CC Endocytosis; Receptor; Transmembrane; EGF-like domain; Repeat;
KW Glycoprotein; LDL; Lipid transport; Cholesterol metabolism; Signal;
FT SIGNAL 1 28
FT PROPEP 29 81
FT CHAIN 82 2215
FT DOMAIN 82 2138
FT TRANSMEM 2139 2159
FT DOMAIN 2160 2215
FT REPEAT 136 147
FT REPEAT 232 243
FT REPEAT 441 452
FT REPEAT 521 532
FT REPEAT 562 573
FT DOMAIN 803 977
FT REPEAT 803 806
FT REPEAT 847 850
FT REPEAT 891 894
FT REPEAT 934 937
FT REPEAT 974 977
FT DOMAIN 1026 1072
FT DOMAIN 1076 1114
FT DOMAIN 1115 1155

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:37:25 ; Search time 50.5714 Seconds
(without alignments)
153.082 Million cell updates/sec

Title: US-09-883-727A-53
Perfect score: 185
Sequence: 1 GCDGFKRLGCTYGFKTDKKGCEAFCTCNT 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriapi.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	45.7	119	5 Q9TWQ7	Q9twq7 haementeria
2	84	45.4	75	5 Q96858	Q96858 hirudo nipp
3	83.5	45.1	120	5 Q25065	Q25065 haementeria
4	81	43.8	59	5 P82107	P82107 hirudo medi
5	76	41.1	2224	5 Q44131	Q44131 caenorhabdi
6	74.5	40.3	119	5 Q9TWX3	Q9twx3 haementeria
7	69.5	37.6	1394	5 Q8MST1	Q8mst1 drosophila
8	69.5	37.6	1511	5 Q9VB21	Q9vb21 drosophila
9	67.5	36.5	1048	13 Q8AWW5	Q8aww5 gallus gall
10	65.5	35.4	1028	11 Q9JLLO	Q9jll0 mus musculu
11	65	35.1	725	10 Q9ZVU2	Q9zvu2 arabidopsis
12	64.5	34.9	1036	4 Q9NZV1	Q9nzv1 homo sapien
13	62.5	33.8	544	13 Q8AVH7	Q8avh7 xenopus lae
14	61.5	33.2	898	5 Q8MQG2	Q8mqg2 caenorhabdi
15	61.5	33.2	909	5 Q17429	Q17429 caenorhabdi
16	61.5	33.2	960	5 Q8MM07	Q8mm07 caenorhabdi

17	60	32.4	80	5 Q9BIV4	Q9biv4 crassostrea
18	60	32.4	599	10 Q9C8A8	Q9c8a8 arabidopsis
19	60	32.4	661	10 Q9C6G7	Q9c6g7 arabidopsis
20	60	32.4	999	4 Q9NQ36	Q9nq36 homo sapien
21	58.5	31.6	327	11 Q8BPM8	Q8bpm8 mus musculu
22	58.5	31.6	550	11 Q9JJZ5	Q9jjz5 mus musculu
23	58.5	31.6	791	6 Q9GK49	Q9gk49 bos taurus
24	58.5	31.6	800	6 Q8SQB9	Q8sqb9 bos taurus
25	57.5	31.1	74	5 Q8WQ15	Q8wq15 ostrea edul
26	57.5	31.1	94	6 Q9SN85	Q9sn85 ovis aries
27	57.5	31.1	112	5 Q8MVK6	Q8mvk6 boitenia v1
28	57.5	31.1	251	5 Q24774	Q24774 enchytraeus
29	57	30.8	1785	13 Q8JHV7	Q8jhv7 brachydanio
30	56.5	30.5	1161	5 Q8T854	Q8t854 dictyosteli
31	56	30.3	76	5 Q8MUZ7	Q8muz7 crassostrea
32	56	30.3	626	4 Q8ND91	Q8nd91 homo sapien
33	56	30.3	969	4 Q9GKG6	Q9gkg6 homo sapien
34	56	30.3	988	4 Q8IWX4	Q8iwx4 homo sapien
35	56	30.3	1007	13 Q8JI28	Q8ji28 xenopus lae
36	55.5	30.0	332	12 Q9W110	Q9w110 bovine herp
37	55.5	30.0	553	4 Q8NBV0	Q8nbv0 homo sapien
38	55.5	30.0	553	4 Q9NZL7	Q9nzl7 homo sapien
39	55.5	30.0	553	4 Q8IUX8	Q8iux8 homo sapien
40	55.5	30.0	554	4 Q9NY67	Q9ny67 homo sapien
41	55.5	30.0	558	4 Q9UFG6	Q9ufk6 homo sapien
42	55.5	30.0	1212	5 Q9XZ29	Q9xz29 drosophila
43	55.5	30.0	1827	13 Q8JHV6	Q8jhv6 brachydanio
44	55	29.7	299	11 Q8BX64	Q8bx64 mus musculu
45	55	29.7	320	10 Q40862	Q40862 picea glauc

ALIGNMENTS

RESULT 1

Q9TWQ7 PRELIMINARY; PRT; 119 AA.
AC Q9TWQ7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Antistatin isoform B, ATs isoform B-BLOOD coagulation factor XA inhibitor.
OS Haementeria officinalis (Mexican leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Rhynchobdellida; Glossiphoniidae; Haementeria.
OX NCBI_TaxID=6410;
RN [1]
RP SEQUENCE.
RX MEDLINE=94097222; PubMed=8271959;
RA Dunwiddie C.T., Waxman L., Vlasuk G.P., Friedman P.A.;
RT "Purification and characterization of inhibitors of blood coagulation factor Xa from hematophagous organisms";
RL Meth. Enzymol. 223:291-312(1993).
DR HSP; P15358; 1SKZ.
DR InterPro: IPR004094; Antistatin.
DR Pfam: PF02822; Antistatin; 2.
SQ SEQUENCE 119 AA; 13426 MW; A08DA51B6CF2E265 CRC64;

Query Match 45.7%; Score 84.5; DB 5; Length 119;
Best Local Similarity 48.1%; Pred. No. 8.8e-05;
Matches 13; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 CDGFKRLGCTYGFKTDKKGCEAFCTC 28
| | | | | : | | | : | | | |
Db 28 CSGVRMCHCPHGFGFSRYGCE-FCKC 53

RESULT 2

Q96858 PRELIMINARY; PRT; 75 AA.
ID Q96858
AC Q96858;
DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Guamerin (Fragment).
 GN GUMI.
 OS Hirudo nipponia (Leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinina; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=42736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95293987; PubMed=7775446;
 RA Jung H.I., Kim S.I., Ha K.S., Joe C.O., Kang K.W.;
 RT "Isolation and characterization of guamerin, a new human leukocyte
 RT elastase inhibitor from Hirudo nipponia.";
 RL J. Biol. Chem. 270:13879-13884(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA KO J.K., Choi K.H., Jung H., Hong S.J., Kim D.R., Lee J.Y., Kim Y.H.,
 RA Kim S.C., Kang K.W.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U38282; AAD09442.1; -;
 DR HSSP; P80302; 1HTA.
 DR InterPro; IPR004094; Antistatin.
 DR InterPro; IPR006210; IEGF.
 DR Pfam; PF02822; Antistatin; 1.
 DR SMART; SM00181; EGF; 1.
 FT NON_TER 1
 SQ SEQUENCE 75 AA; 8081 MW; E81CA7D9FF17D53F CRC64;

 Query Match 45.48; Score 84; DB 5; Length 75;
 Best Local Similarity 48.18; Pred. No. 6.6e-05;
 Matches 13; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

 Qy 2 CDGFKRLGCTGYGKTDKKGCEAFCTC 28
 | : : : | | | : | | | | |
 Db 35 CTARCMIFCPNGFKVDNGCEYPTC 61

 RESULT 3
 Q25065
 ID Q25065 PRELIMINARY; PRT; 120 AA.
 AC Q25065;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Ghilanten.
 OS Haementeria ghiliani (Amazon leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinina; Hirudinea;
 OC Rhynchobdellida; Glossiphoniidae; Haementeria.
 OX NCBI_TaxID=6409;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96363468; PubMed=8746634;
 RA Brankamp R.G., Sreekrishna K., Smith P.L., Blankenship D.T.,
 RA Cardin A.D.;
 RT "Expression of a synthetic gene encoding the anticoagulant-
 RT antimeastatic protein ghilanten by the methylotropic yeast Pichia
 RT pastoris.";
 RL Protein Expr. Purif. 6:813-820(1995).
 DR EMBL; U20787; AAA5645.1; -;
 DR HSSP; P15358; ISK2.
 DR InterPro; IPR004094; Antistatin.
 DR Pfam; PF02822; Antistatin; 2.
 SQ SEQUENCE 120 AA; 13409 MW; A10BBE519E776E97 CRC64;

 Query Match 45.18; Score 83.5; DB 5; Length 120;
 Best Local Similarity 44.88; Pred. No. 0.00012;
 Matches 13; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

 Qy 2 CDGFKRLGCTGYGKTDKKGCEAFCTC 30
 | : : : | | | : | | | | |
 Db 29 CSGVRCVYCSHGFSRQYCEV-CRCRT 56

RESULT 4
 P82107
 ID P82107 PRELIMINARY; PRT; 59 AA.
 AC P82107;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE BELLASTASIN (Bdellin A).
 OS Hirudo medicinalis (Medicinal leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinina; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=98237591; PubMed=9578479;
 RA Moser M., Auerwald E., Mentele R., Eckerskorn C., Fritz H., Fink E.;
 RT "Bdellastatin, a serine protease inhibitor of the antistatin family
 RT from the medicinal leech (Hirudo medicinalis). Primary structure,
 RT expression in yeast, and characterization of native and recombinant
 RT inhibitor.";
 RL Eur. J. Biochem. 253:212-220(1998).
 CC -1- FUNCTION: STRONG INHIBITOR OF MAMMALIAN TRYPSIN, PLASMIN AND
 CC ACROSIN.
 CC -1- MASS SPECTROMETRY: MW=6332.6; METHOD=ELECTROSPRAY.
 CC -1- SIMILARITY: BELONGS TO THE ANTISTATIN FAMILY.
 DR HSSP; P80302; 1HTA.
 DR InterPro; IPR004094; Antistatin.
 DR Pfam; PF02822; Antistatin; 1.
 KW Serine protease inhibitor.
 FT ACT_SITE 34 35 REACTIVE BOND (BY SIMILARITY).
 FT DISULFID 10 21 BY SIMILARITY.
 FT DISULFID 15 26 BY SIMILARITY.
 FT DISULFID 28 48 BY SIMILARITY.
 FT DISULFID 33 52 BY SIMILARITY.
 FT DISULFID 37 54 BY SIMILARITY.
 SQ SEQUENCE 59 AA; 6343 MW; 43BA5BE2D0E403A9 CRC64;

 Query Match 43.88; Score 81; DB 5; Length 59;
 Best Local Similarity 44.48; Pred. No. 0.00014;
 Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

 Qy 2 CDGFKRLGCTGYGKTDKKGCEAFCTC 28
 | : : : | | | : | | | | |
 Db 28 CSDLHCKVKCEHGFKKDDNGCEYACIC 54

 RESULT 5
 O44131
 ID O44131 PRELIMINARY; PRT; 2224 AA.
 AC O44131;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hypothetical protein C08G9.2.
 GN C08G9.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Geisel C., Stellyes L.;
 RT "The sequence of C. elegans cosmid C08G9.";

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;

RA Waterston R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF036687; AAB88311.2; -

DR HSSP; P10646; IADZ.

DR WormPep; C08G9.2; CE24793.

DR InterPro; IPR004094; Antistatin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR002223; Kunitz_BPTI.

DR InterPro; IPR000716; Thyroglobulin_1.

DR InterPro; IPR002221; WAP.

DR InterPro; IPR007087; Znf.C2H2.

DR Pfam; PF02822; Antistatin; 6.

DR Pfam; PF00014; Kunitz_BPTI; 1.

DR Pfam; PF00086; Thyroglobulin_1; 6.

DR Pfam; PF00095; wap; 9.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz_BPTI; 1.

DR ProDom; PD001224; WAP; 4.

DR SMART; SM00131; KU; 1.

DR SMART; SM00211; TY; 6.

DR SMART; SM00217; WAP; 9.

DR PROSITE; PS00317; 4_DISULFIDE_CORE; 9.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS00484; THYROGLOBULIN_1; 4.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.

KW Hypothetical protein; Protease inhibitor; Serine protease inhibitor.

SQ SEQUENCE 2224 AA; 242101 MW; 46C805043A971660 CRC64;

Query Match 41.1%; Score 76; DB 5; Length 2224;

Best Local Similarity 50.0%; Pred. No. 0.028;

Matches 14; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

QY 2 CDGFKC-RLGCTYGFKTKKGCFAFCTC 28

DB 905 CEKLEGTGCEYGHKKARGC-ATCDC 931

RESULT 6

Q9TWX3

ID Q9TWX3

AC Q9TWX3

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Ghilanten.

OS Haementeria ghiliani (Amazon leech).

OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;

OC Rhynchobdellida; Glossiphoniidae; Haementeria.

OX NCBI_TaxID=6409;

RN [1]

RP SEQUENCE.

RX MEDLINE-92126887; PubMed-1772984;

RA Brankamp R.G., Mauley G.G., Blankenship D.T., Bowlin T.L.,

Cardin A.D.;

RT "Studies on the anticoagulant, antimetastatic and heparin-binding

RT properties of ghilanten-related inhibitors.";

RL Blood Coagul. Fibrinolysis 2:161-166(1991).

DR HSSP; P15358; 1SKZ.

DR InterPro; IPR004094; Antistatin.

DR Pfam; PF02822; Antistatin; 2.

SQ SEQUENCE 119 AA; 13318 MW; BE94805DAF85F82 CRC64;

Query Match 40.3%; Score 74.5; DB 5; Length 119;

Best Local Similarity 41.4%; Pred. No. 0.0026;

Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 2 CDGFKCRLGCTYGFKTKKGCFAFCTCNT 30

DB 28 CPEVRCRVYCSHGFRSRYGCEV-CRCRT 55

RESULT 7

Q8MST1

ID Q8MST1

AC Q8MST1

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Ip117788p.

GN CG5639.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY118620; AAM4989.1; -

DR FlyBase; FBgn0039527; CG5639.

DR InterPro; IPR004094; Antistatin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000716; Thyroglobulin_1.

DR InterPro; IPR002221; WAP.

DR Pfam; PF02822; Antistatin; 3.

DR Pfam; PF00086; Thyroglobulin_1; 3.

DR Pfam; PF00095; wap; 2.

DR PRINTS; PR00003; 4DISULFHCORE.

DR SMART; SM00211; TY; 4.

DR SMART; SM00217; WAP; 2.

DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.

DR PROSITE; PS00022; EGF_1; 3.

DR PROSITE; PS00484; THYROGLOBULIN_1; 3.

DR PROSITE; PS00678; WD_REPEATS_1; 1.

SQ SEQUENCE 1394 AA; 152773 MW; EDE4F62E8DAE5D8F CRC64;

Query Match 37.6%; Score 69.5; DB 5; Length 1394;

Best Local Similarity 46.4%; Pred. No. 0.16;

Matches 13; Conservative 1; Mismatches 13; Indels 1; Gaps 1;

QY 2 CDGFKCRLGCTYGFKTKKGCFAFCTCN 29

DB 256 CDTNCAAYCEYGFKNHNGCPT-CBCS 282

RESULT 8

Q9VB21

ID Q9VB21

AC Q9VB21

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE CG5639 protein.

GN CG5639.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RESULT 9					
Q8AWW5	QBAAWS	PRELIMINARY;	PRT;	1048 AA.	
AC	QBAWW5;				
DT	01-MAR-2003 (TReMBLrel. 23, Created)				
DT	01-MAR-2003 (TReMBLrel. 23, Last sequence update)				
DE	01-MAR-2003 (TReMBLrel. 23, Last annotation update)				
DE	Cysteine-rich motorneuron 1.				
GN	CRIM1.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCB1_TaxID=9031;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22395914; PubMed=12508231;				
RA	Kolle G., Jansen A., Yamada T., Little M.;				
RT	"In ovo electroporation of Criml in the developing chick spinal				
RT	cord.";				
RL	Dev. Dyn. 226:107-111(2003).				
DR	EMBL; AY098584; AAM28339.1;				
SQ	SEQUENCE 1048 AA; 114942 MW; 25E4D82C40B08231 CRC64;				
Query Match 36.5%; Score 67.5; DB 13; Length 1048;					
Best Local Similarity 44.8%; Pred. No. 0.23;					
Matches 13; Conservative 3; Mismatches 10; Indels 3; Gaps					
QY	2 CDGFF-KCRIGCTYGFKTKKGCEAFCTC 28				
	: :				
Db	517 CTGLISGCSLDSCSFGQTDAHNCE-ICQC 544				
RESULT 10					
Q9JLL0	Q9JLL0	PRELIMINARY;	PRT;	1028 AA.	
AC	Q9JLL0;				
DT	01-OCT-2000 (TReMBLrel. 15, Created)				
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)				
DT	01-OCT-2002 (TReMBLrel. 22, Last annotation update)				
DE	Cysteine-rich repeat-containing protein CRIM1 precursor				
DE	(Fragment).				
GN	CRIM1				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20108580; PubMed=10642437;				
RA	Kolle G., Georgasi K., Holmes G.P., Little M.H., Yamada T.;				
RT	*CRIM1, a novel gene encoding a cysteine-rich repeat protein, is				
RT	developmentally regulated and implicated in vertebrate CNS development				
RT	and organogenesis.";				
RL	Mech. Dev. 90:181-193(2000).				
DR	EMBL; AF168680; AAF34410.1; -				
DR	HSSP; PI5358; ISKZ.				
DR	MGD; MGI:1354756; Criml.				
DR	InterPro; IPR004094; Antistatin.				
DR	InterPro; IPR000867; Insl_gro_fac_pr.				
DR	InterPro; IPR001007; VWF.C.				
DR	Pfam; PF02822; Antistatin; 4.				
DR	Pfam; PR00093; vwc; 6.				
DR	SMART; SM00121; IB; 1.				
DR	SMART; SM00214; VWC; 6.				
DR	PROSITE; PS01208; VWF_C; 6.				
KW	Signal.				
FT	NON_TER 1				
FT	SIGNAL <1 25				
SQ	SEQUENCE 1028 AA; 113160 MW; 743058AA481D5ED8 CRC64;				
Query Match 35.4%; Score 65.5; DB 11; Length 1028;					

Best Local Similarity 48.3%; Pred. No. 0.45; 11; Indels 3; Gaps 2;
Matches 14; Conservative 1; Mismatches 11; Indels 3; Gaps 2;

QY 2 CDGFK--CRIGCTYGFKTDKKGCEAFCTC 28
Db 496 CLGLKRACTLDGPFGLTDVHNC-ICQC 523

RESULT 11
Q9ZVU2 ID Q9ZVU2 PRELIMINARY; PRT; 725 AA.
AC Q9ZVU2
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE T5A14.16 protein.
GN T5A14.16
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altati H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005223; AAD10653.1; -
DR InterPro: IPR004146; DC1.
DR InterPro: IPR001965; Znf_PHD.
DR Pfam: PF03107; DC1; 4.
DR SMART: SM00249; PHD; 2.
SQ SEQUENCE 725 AA; 82501 MW; B6EFB4D7CC4B447 CRC64;

Query Match 35.1%; Score 65; DB 10; Length 725;
Best Local Similarity 44.4%; Pred. No. 0.38;
Matches 12; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 3 DGFKC---RLGCTYGFKTDKKGCEAF 25
Db 468 DIFKICGGRIGCGFYKCKDEKNCDEF 494

RESULT 12
Q9NZV1 ID Q9NZV1 PRELIMINARY; PRT; 1036 AA.
AC Q9NZV1; Q9H318;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cysteine-rich repeat-containing protein S52 precursor (CRIM1 protein).
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20108580; PubMed=10642437;
RA Kelle G., Georgas K., Holmes G.P., Little M.H., Yamada T.;
RT "CRIM1, a novel gene encoding a cysteine-rich repeat protein, is developmentally regulated and implicated in vertebrate CNS development and organogenesis.";
RT Mech. Dev. 90:181-193(2000).
RN [2]
RP SEQUENCE OF 112-1036 FROM N.A.
RA Kelle G., Georgas K., Holmes G.P., Little M.H., Yamada T.;
RT "CRIM1, a novel gene encoding a cysteine-rich repeat protein, is developmentally regulated and implicated in vertebrate CNS development and organogenesis.";

Mech. Dev. 0:0-0(2000).
DR EMBL: AF167706; AAF34409.1; -
DR EMBL: AF168681; AAG37011.1; -
DR HSP: P15358; 1SK2.
DR Genew: HGNC:2359; CRIM1.
DR InterPro: IPR004094; Antistatin.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000169; Shprot_acsite.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF02822; Antistatin; 4.
DR Pfam: PF00093; vwc; 6.
DR SMART: SM00121; IB; 1.
DR SMART: SM00214; VMC; 6.
DR PROSITE: PS00139; THIOLEPROTEASE_CYS; 1.
DR PROSITE: PS01208; WIFC; 6.
KW Signal.
FT SIGNAL. 1 34 POTENTIAL.
SQ SEQUENCE 1036 AA; 113737 MW; 10CBF02A5C579C27 CRC64;

Query Match 34.9%; Score 64.5; DB 4; Length 1036;
Best Local Similarity 50.0%; Pred. No. 0.64;
Matches 11; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 7 CRIGCTYGFKTDKKGCEAFCTC 28
Db 512 CTLNCPFGFLTDACNCE-ICEC 532

RESULT 13
Q8AVH7 ID Q8AVH7 PRELIMINARY; PRT; 544 AA.
AC Q8AVH7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to EGF-like-domain, multiple 6.
OS Xenopus laevis (African clawed frog).
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Euteleostomi;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC042275; AAH42275.1; -
SQ SEQUENCE 544 AA; 61048 MW; 24865A62D2EAB7B5 CRC64;

Query Match 33.8%; Score 62.5; DB 13; Length 544;
Best Local Similarity 52.0%; Pred. No. 0.66;
Matches 13; Conservative 3; Mismatches 4; Indels 5; Gaps 2;

QY 7 CRIG----CTYGFKTDKKG-CEAFC 26
Db 47 CRYGKAECCYGNKRNKGCEAVC 71

RESULT 14
Q8MQG2 ID Q8MQG2 PRELIMINARY; PRT; 898 AA.
AC Q8MQG2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE B0024.14d protein.
GN B0024.14 OR B0024.14D.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

```

QY      6 KRLRGCTGYGFTDKKGCBAFTC 28
      : | | | | | : | | : |
Db      396 QCHKCLYGFETNSAGC-SLKC 417

Search completed: August 5, 2003, 09:40:57
Job time : 52.5714 secs

```

15
16

```

RESULT 15
Q17429
ID Q17429 PRELIMINARY; PRT; 909 AA.
AC Q17429;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE B0024.14 protein.
DE B0024.14
GN B0024.14
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN RN SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN RN SEQUENCE FROM N.A.
RX MEDLINE=95069613; PubMed=9851916;
none;
RA "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RT Science 282:2012-2018(1998).

```

1:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:37:24 ; Search time 15 Seconds
(without alignments)
64.112 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYEYE 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	74.6	511	2 S24345	Balbani ring 1 pr
2	43	72.9	749	2 A45294	Balbani ring 2.1
3	43	72.9	1088	1 P1XRPR	inner layer protei
4	43	72.9	1088	2 S39261	VPI protein - porc
5	42	71.2	149	2 F69876	conserved hypotet
6	42	71.2	790	2 T19683	hypothetical prote
7	41	69.5	218	2 T36242	probable precorrin
8	41	69.5	358	2 G70601	hypothetical prote
9	40	67.8	222	2 G97704	hypothetical prote
10	40	67.8	259	2 F90502	hypothetical prote
11	40	67.8	500	2 B83910	hypothetical prote
12	40	67.8	2467	2 D71437	hypothetical prote
13	39	66.1	125	2 AF1134	probable resistanc
14	39	66.1	144	2 G69456	probable secreted
15	39	66.1	207	2 T10142	hypothetical prote
16	39	66.1	347	2 T24921	modA.1 protein - p
17	39	66.1	1123	2 T22608	hypothetical prote
18	39	66.1	1687	2 T43144	hypothetical prote
19	39	66.1	3396	2 T26113	vitellogenin II pr
20	38	64.4	203	2 E70150	hypothetical prote
21	38	64.4	249	2 A81300	hypothetical prote
22	38	64.4	285	2 E40590	probable diaminopi
23	38	64.4	536	2 F89996	motA homolog lair
24	38	64.4	617	2 T16189	hypothetical prote
25	38	64.4	650	2 JC4673	hypothetical prote
26	38	64.4	929	2 T34206	protein kinase (EC
27	38	64.4	1196	2 JQ1467	hypothetical prote
28	38	64.4	1196	2 T46430	toxin, nontoxic co
29	38	64.4	2899	2 T21546	botulinum neurotox
					hypothetical prote

30 38 64.4 2915 2 G87867 protein F36A2.13 (

31 38 64.4 6831 2 A88852 protein unc-22 lim

32 38 64.4 6839 2 S57242 twitchin [similar

33 38 64.4 7160 2 T27935 hypothetical prote

34 37 62.7 82 2 F64308 hypothetical prote

35 37 62.7 219 2 F71710 hypothetical prote

36 37 62.7 291 2 D81677 conserved hypotet

37 37 62.7 301 2 S62087 hrpF protein - pse

38 37 62.7 305 2 T24283 hypothetical prote

39 37 62.7 369 2 B71624 hypothetical prote

40 37 62.7 372 2 T24392 hypothetical prote

41 37 62.7 424 2 A88979 protein F37B4.7 [i

42 37 62.7 535 2 T29677 hypothetical prote

43 37 62.7 870 2 T47454 lipoxigenase AtLOX

44 37 62.7 896 2 JQ2391 lipoxygenase (EC 1

45 37 62.7 1084 2 C82931 hypothetical prote

ALIGNMENTS

RESULT 1

S24345

Balbani ring 1 protein - midge (Chironomus tentans) (fragment)

C;Species: Chironomus tentans

C;Date: 13-Jan-1995 #sequence_revision 04-Mar-2000 #text_change 21-Jul-2000

C;Accession: S24345

R;Paulson, G.; Hoeeg, C.; Bernholm, K.; Wieslander, L.

J. Mol. Biol. 225, 349-361, 1992

A;Title: Balbani ring 1 gene in Chironomus tentans. Sequence organization and dynamics

A;Reference number: S24345; MUID:9227646; PMID:1593624

A;Accession: S24345

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-511 <PAU>

A;Cross-references: EMBL:X64322; NID:g7041; PIDN:CAA45607.1; PID:g7042

C;Superfamily: unassigned Balbani ring proteins

Query Match 74.6%; Score 44; DB 2; Length 511;

Best Local Similarity 77.8%; Pred. No. 7.9;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEEYEYE 10

Db 111 DEYEYE 119

RESULT 2

A45294

Balbani ring 2.1 - midge (Chironomus tentans) (fragment)

C;Species: Chironomus tentans

C;Date: 25-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C;Accession: A45294

R;Wieslander, L.; Paulson, G.

Proc. Natl. Acad. Sci. U.S.A. 89, 4578-4582, 1992

A;Title: Sequence organization of the Balbani ring 2.1 gene in Chironomus tentans.

A;Reference number: A45294; MUID:92262483; PMID:1584794

A;Accession: A45294

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-749 <WIE>

A;Cross-references: GB:M89909; NID:gi56605; PIDN:AAA28264.1; PID:gi56606

A;Note: sequence extracted from NCBI backbone (NCBIN:102269, NCBIN:102271, NCBIN:102273,

C;Superfamily: unassigned Balbani ring proteins

C;Keywords: tandem repeat

Query Match 72.9%; Score 43; DB 2; Length 749;

Best Local Similarity 87.5%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EEEYEYE 10

Db 122 DEYEYE 129

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: A69580; MUID:98044033; PMID:9384377
 A;Accession: F69876
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-149 <KUN>
 A;Cross-references: GB:299112; GB:AL009126; NID:g26333902; PIDN:CAB13413.1; PID:e1185131
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: ylmF

Query Match 71.2%; Score 42; DB 2; Length 149;
 Best Local Similarity 87.5%; Pred. No. 4.2; Mismatches 1; Indels 0; Gaps 0;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NEEYEY 9
 Db 13 DEEY 20

RESULT 6
 T19683
 hypothetical protein C33D9.6 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T19683
 R;Lloyd, C.
 submitted to the EMBL Data Library, December 1995
 A;Reference number: Z19162
 A;Accession: T19683
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-790 <WIL>
 A;Cross-references: EMBL:Z68159; PIDN:CRA92287.1; GSPDB:GN00022; CESP:C33D9.6
 C;Genetics:
 A;Gene: CESP:C33D9.6
 A;Map position: 4
 A;Introns: 117/3; 254/1; 352/1; 371/1; 436/3; 477/1; 558/3; 665/1; 724/1
 Query Match 71.2%; Score 42; DB 2; Length 790;
 Best Local Similarity 77.8%; Pred. No. 28; Mismatches 7; Conservative 1; Indels 1; Gaps 0;
 Qy 1 PNEEY 9
 Db 420 PEEY 428

RESULT 7
 T36242
 probable precorrin-8X methylmutase - *Streptomyces coelicolor*
 C;Species: *Streptomyces coelicolor*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C;Accession: T36242
 R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, March 1999
 A;Reference number: Z21577
 A;Accession: T36242
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-218 <OLI>
 A;Cross-references: EMBL:AL049573; PIDN:CAB40340.1; GSPDB:GN00070; SCOEDB:SCE39.32
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SCE39.32
 C;Superfamily: *Methanobacterium thermoautotrophicum* precorrin isomerase
 Query Match 69.5%; Score 41; DB 2; Length 218;

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: A69580; MUID:98044033; PMID:9384377
 A;Accession: F69876
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-149 <KUN>
 A;Cross-references: GB:299112; GB:AL009126; NID:g26333902; PIDN:CAB13413.1; PID:e1185131
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: ylmF

Query Match 72.9%; Score 43; DB 1; Length 1088;
 Best Local Similarity 77.8%; Pred. No. 28; Mismatches 1; Indels 0; Gaps 0;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NEEYEY 10
 Db 186 NEKYRYEY 194

RESULT 4
 S39261
 VPI protein - porcine rotavirus
 C;Species: porcine rotavirus
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
 C;Accession: S39261
 R;Almanza, L.; Arias, C.F.; Lopez, S.
 submitted to the EMBL Data Library, November 1993
 A;Description: Amino acid sequence of the Porcine rotavirus YM VPI protein.
 A;Reference number: S39261
 A;Accession: S39261
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1088 <ALM>
 A;Cross-references: EMBL:X76486; NID:9434312; PIDN:CAA54024.1; PID:g434313
 C;Superfamily: rotavirus inner layer protein VPI
 Query Match 72.9%; Score 43; DB 2; Length 1088;
 Best Local Similarity 77.8%; Pred. No. 28; Mismatches 1; Indels 0; Gaps 0;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NEEYEY 10
 Db 186 NEKYRYEY 194

RESULT 5
 F69876
 conserved hypothetical protein ylmF - *Bacillus subtilis*
 C;Species: *Bacillus subtilis*
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C;Accession: F69876
 R;Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 A.; Ehrlich, S.D.; Emmeron, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallex
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mueel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivcita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 Query Match 72.9%; Score 43; DB 2; Length 1088;
 Best Local Similarity 77.8%; Pred. No. 28; Mismatches 1; Indels 0; Gaps 0;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NEEYEY 10
 Db 186 NEKYRYEY 194

inner layer protein VPI - porcine rotavirus C (strain Gottfried)
 N;Alternate names: core protein VPI
 C;Species: porcine rotavirus C
 C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 08-Apr-1994
 C;Accession: A33749
 R;Fukuhara, N.; Nishikawa, K.; Gorziglia, M.; Kapikian, A.Z.
 Virology 173, 743-749, 1989
 A;Title: Nucleotide sequence of gene segment 1 of a porcine rotavirus strain.
 A;Reference number: A33749; MUID:90085826; PMID:2556853
 A;Accession: A33749
 A;Molecule type: genomic RNA
 A;Residues: 1-1088 <FUK>
 C;Genetics:
 A;Map position: segment 1
 C;Superfamily: rotavirus inner layer protein VPI
 C;Keywords: core protein; glycoprotein
 F:68,236,258,547,760,1048,1072/Binding site: carbohydrate (Asn) (covalent) #status predi

Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EYEYEE 10
|||||||
Db 9 EYEYEE 15

RESULT 8
G70601
hypothetical protein Rv0996 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C;Accession: G70601
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70601
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <COL>
A;Cross-references: GB:Z94752; GB:AL123456; NID:93261731; PIDN:CAB08146.1; PID:G2052130
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0996
C;Superfamily: Mycobacterium leprae probable membrane protein MLCB373.37

Query Match 69.5%; Score 41; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EYEYEE 10
|||||||
Db 181 EYEYEE 187

RESULT 9
G97704
hypothetical protein RC0039 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: G97704
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: G97704
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL02577.1; PID:G15619074; GSPDB:GNO0173
C;Genetics:
A;Gene: RC0039
C;Superfamily: Rickettsia prowazekii hypothetical protein RP030

Query Match 67.8%; Score 40; DB 2; Length 222;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NEEYEEYE 10
|||||||
Db 45 NEEFEYEME 53

RESULT 10
F90502
hypothetical protein SSO3179 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: F90502
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, I
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-259 <KUR>
A;Cross-references: GB:AE006641; NID:G13816612; PIDN:AAK43277.1; GSPDB:GNO0155
C;Genetics:
A;Gene: SSO3179

Query Match 67.8%; Score 40; DB 2; Length 259;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 9
|||||||
Db 200 PNLDDYDEY 208

RESULT 11
B83910
hypothetical protein BH2082 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83910
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-500 <STO>
A;Cross-references: GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA05801.1; GSPDB:GNO
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2082

Query Match 67.8%; Score 40; DB 2; Length 500;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
|||||||
Db 128 PGEEFTYEQ 137

RESULT 12
D71437
probable resistance gene - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: Columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002
C;Accession: D71437
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dir
P.; Wedler, H.; Wedler, B.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvas, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
C.; Chalwatzi, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: D71437
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2467 <BEV>
A;Cross-references: GB:Z97342; NID:G2245031; PID:G2245050

C;Genetics:

A;Map position: 4COP9-4G3845

Query Match 67.8%; Score 40; DB 2; Length 2467;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 BEVEYEYE 10

DB 2136 DEYEYED 2143

RESULT 13

AF1134

Probable secreted protein lmo0477 [imported] - Listeria monocytogenes (strain EGD-e)

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: AF1134

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Klett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative Genomics of Listeria species.

A;Reference number: AB1077; MUID: 21537279; PMID:11679669

A;Accession: AF1134

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-125 <GLA>

A;Cross-references: GB:NC 003210; PIDN:CAC98556.1; PID:g16409853; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo0477

C;Superfamily: Bacillus thuringiensis plasmid p612 hypothetical protein 3

Query Match

Best Local Similarity 66.1%; Score 39; DB 2; Length 125;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10

DB 75 PNLKYEYSYD 84

RESULT 14

G69456

hypothetical protein AF1656 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 04-Mar-2000

C;Accession: G69456

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: G69456

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-144 <KLE>

A;Cross-references: GB:AE000989; GB:AE000782; NID:g2689312; PIDN:AB89614.1; PID:g264892

C;Superfamily: Archaeoglobus fulgidus hypothetical protein AF1656

Query Match

Best Local Similarity 66.1%; Score 39; DB 2; Length 144;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 8

DB 64 PNEEYSYD 71

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:35:59 ; Search time 11 Seconds
(without alignments)
42.752 Million cell updates/sec

Title: US-09-883-727A-125
Perfect score: 59
Sequence: 1 PNEYEYEYE 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	72.9	1088	1 RRPO_ROTGP	P17699 porcine rot
2	43	72.9	1737	1 CO4_EAT	P08649 rattus norv
3	42	71.2	435	1 ERFL_TETH	Q08649 rattus norv
4	39	66.1	144	1 YG56_ARCFU	Q08617 archaerhena
5	39	66.1	207	1 MODA_BPT4	P39423 bacterioph
6	39	66.1	1111	1 UN84_CABEL	Q20745 caenorhabd
7	39	66.1	1687	1 VIT2_FUNHE	Q98893 fundulus he
8	38	64.4	249	1 DAPF_CAMJE	Q9pm48 campylobact
9	38	64.4	285	1 LAFT_VIBPA	Q03477 vibrio para
10	38	64.4	1196	1 EXCN_CLOBO	P46081 clostridium
11	37	62.7	82	1 Y070_METJA	Q60373 methanococc
12	37	62.7	289	1 Y665_CHLMU	Q9pk08 chlamydia m
13	37	62.7	409	1 HEMI_TREVO	Q97b68 thermoplas
14	37	62.7	424	1 FLOI_CABEL	Q45166 caenorhabd
15	37	62.7	535	1 FLOI_CABEL	Q22931 caenorhabd
16	37	62.7	896	1 LOXC_ARATH	P38418 arabidopsis
17	37	62.7	1088	1 RRPO_ROTBR	P17468 bovine rota
18	37	62.7	1088	1 RRPO_ROTBU	P21615 bovine rota
19	37	62.7	1208	1 PCPI_SCHPO	Q92351 schizosacch
20	37	62.7	1448	1 PK3G_HUMAN	O75747 homo sapien
21	37	62.7	1894	1 BGS2_SCHPO	O13967 schizosacch
22	37	62.7	1955	1 BGS4_SCHPO	O74475 schizosacch
23	36	61.0	222	1 YP32_YEAST	P51996 saccharomyc
24	36	61.0	263	1 KLR_PRANA	P32824 praomys nat
25	36	61.0	273	1 DAPF_HELPJ	Q921x5 helicobacte
26	36	61.0	273	1 DAPF_HELPJ	O25290 helicobacte
27	36	61.0	296	1 CYOA_BUCAI	P57544 buchnera ap
28	36	61.0	308	1 PAP2_FOWPV	P15916 fowlpox vir
29	36	61.0	333	1 PAP2_VACCC	P21033 vaccinia vi
30	36	61.0	333	1 PAP2_VACCV	P07617 vaccinia vi
31	36	61.0	333	1 PAP2_VARV	P33052 variola vir
32	36	61.0	394	1 VORA_PYRFU	Q51801 pyrococcus
33	36	61.0	416	1 MYC_AVIOK	P12523 avian retro

34	36	61.0	416	1 MYC_CHICK	P01109 gallus gall
35	36	61.0	419	1 MYC1_XENLA	P06171 xenopus lae
36	36	61.0	420	1 MYC2_XENLA	P15171 xenopus lae
37	36	61.0	421	1 MYC_AVIM2	P10395 avian myelo
38	36	61.0	422	1 MYC_AVIMC	P01110 avian myelo
39	36	61.0	422	1 MYC_AVIMD	P06295 avian myelo
40	36	61.0	423	1 MYC_AVIME	P06647 avian retro
41	36	61.0	423	1 OMD_MOUSE	O35103 mus musculu
42	36	61.0	454	1 BLHI_YEAST	Q01532 saccharomyc
43	36	61.0	573	1 ILVI_HABIN	P45261 haemophilus
44	36	61.0	831	1 RPOP_GELSP	O03685 gelatinospo
45	36	61.0	941	1 CHR1_XENLA	Q91713 xenopus lae

ALIGNMENTS

RESULT 1
RRPO_ROTGP STANDARD; PRT; 1088 AA.
ID RRPO_ROTGP
AC P17699;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA-directed RNA polymerase subunit VP1 (EC 2.7.7.48) (Inner layer protein VP1) (Core protein VP1).
GN SI.
OS Porcine rotavirus (strain Gertfried).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OC NCBI_TaxID=10917;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085826; PubMed=2556853;
RA Rukhara N., Nishikawa K., Corziglia M., Kapikian A.Z.;
RT "Nucleotide sequence of gene segment 1 of a porcine rotavirus strain.";
RL Virology 173:743-749(1989).
CC -!- FUNCTION: RNA POLYMERASE ACTIVITY IS THOUGHT TO BE ASSOCIATED WITH VP1.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
CC -!- SUBCELLULAR LOCATION: INNER LAYER.
CC -!- SIMILARITY: BELONGS TO THE ROTAVIRUS VP1 PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M32805; AAA33833.1; --
PIR; A33749; P1XRPR.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007097; RNA_pol_rec.
KW Core protein; RNA-directed RNA polymerase; Transferase.
SQ SEQUENCE 1088 AA; 124954 MW; 7B21BB4B55B0C52E CRC64;

Query Match 72.9%; Score 43; DB 1; Length 1088;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NEEYEYE 10
||:| ||||
Db 186 NEKYRYEYE 194

RESULT 2
CO4_RAT STANDARD; PRT; 1737 AA.
ID CO4_RAT
AC P08649; Q62895; Q8R403;
DT 01-AUG-1988 (Rel. 08, Create:)

DR 15-SEP-2003 (Rel. 42, Last sequence update)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C4 precursor [Contains: C4a anaphylatoxin].
GN C4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Chen C.-B., Wallis R.;
RT "Substrate recognition by zymogen and activated forms of mannose-
RT binding protein-associated serine proteases.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-266 FROM N.A.
RC STRAIN=Brown Norway;
RX MEDLINE=22131572; PubMed=12136338;
RA Walter L., Hurt P., Himmelbauer H., Sudbrak R., Guenther E.;
RT "Physical mapping of the major histocompatibility complex class II and
RT class III regions of the rat.";
RL Immunogenetics 54:268-275(2002).
RN [3]
RP SEQUENCE OF 1656-1737 FROM N.A., AND INDUCTION.
RC STRAIN=Sprague-Dawley; TISSUE=Hepatic stellate cells;
RX MEDLINE=96399092; PubMed=8805663;
RA Fimmel C.-J., Brown K.E., O'Neill R., Kladney R.D.;
RT "Complement C4 protein expression by rat hepatic stellate cells.";
RL J. Immunol. 157:2601-2609(1996).
RN [4]
RP SEQUENCE OF 678-753.
RX MEDLINE=88334568; PubMed=3262196;
RA Cui L.-X., Ferreri K., Hugli T.E.;
RT "Structural characterization of the C4a anaphylatoxin from rat.";
RL Mol. Immunol. 25:663-671(1988).
RN [5]
RP SEQUENCE OF 678-753.
RA Cui L.-X., Ferreri K., Hugli T.E.;
RT "Characterization of rat anaphylatoxins C4a and C5a.";
RL Fed. Proc. 44:991-991(1985).
CC -!- FUNCTION: C4 plays a central role in the activation of the
CC classical pathway of the complement system. It is processed by
CC activated C1 which removes from the alpha chain the C4a
CC anaphylatoxin.
CC -!- FUNCTION: Derived from proteolytic degradation of complement C4,
CC C4a anaphylatoxin is a mediator of local inflammatory processes.
CC It induces the contraction of smooth muscle, increases vascular
CC permeability and causes histamine release from mast cells and
CC basophilic leukocytes.
CC -!- SUBUNIT: This protein is synthesized as a single-chain precursor
CC and, prior to secretion, is enzymatically cleaved to form a trimer
CC of nonidentical chains (alpha, beta and gamma).
CC -!- INDUCTION: Induced in hepatic stellate cells by iron overload and
CC by gamma-interferon.
CC -!- MISCELLANEOUS: C4 is a major histocompatibility complex class-III
CC protein.
CC -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC -!- SIMILARITY: Contains 1 NTR domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 1721.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AY149995; AAM72415.1; -.
DR EMBL; AY091787; AAM14719.1; -.

DR EMBL; U42719; AAA91231.1; ALT_FRAME.
DR PIR; JLO036; JLO036.
DR HSP; P01031; IKJS.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0003811; P:complement activity; NAS.
DR GO; GO:0006958; P:complement activation, classical pathway; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR GO; GO:0045987; P:positive regulation of smooth muscle contra. .; NAS.
DR InterPro; IPR002890; A2M_N.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR001599; MacroglobulinA2.
DR InterPro; IPR001134; Netrin_C.
DR Pfam; PF01821; ANATO; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR ProDom; PD003284; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00477; ALPHA_2 MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW Complement pathway; Glycoprotein; Sulfation; Signal;
KW Inflammatory response; Thioester bond.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 673 COMPLEMENT C4, BETA CHAIN (BY
FT SIMILARITY).
FT PROPEP 674 677 BY SIMILARITY.
FT CHAIN 678 1442 COMPLEMENT C4, ALPHA CHAIN (BY
FT SIMILARITY).
FT PROPEP 1443 1446 BY SIMILARITY.
FT CHAIN 1447 1737 COMPLEMENT C4, GAMMA CHAIN (BY
FT SIMILARITY).
FT PEPTIDE 678 753 C4a ANAPHYLATOXIN.
FT DOMAIN 700 734 ANAPHYLATOXIN-LINE.
FT DISULFID 1583 1735 NTR.
FT DISULFID 700 726 BY SIMILARITY.
FT DISULFID 701 733 BY SIMILARITY.
FT DISULFID 714 734 BY SIMILARITY.
FT CROSSLINK 1005 1008 Iso-glutamyl cysteine thioester (Cys-Gln
FT (By similarity)).
FT MOD_RES 1412 1412 SULFATION (BY SIMILARITY).
FT MOD_RES 1414 1414 SULFATION (POTENTIAL).
FT MOD_RES 1416 1416 SULFATION (BY SIMILARITY).
FT MOD_RES 1676 1676 SULFATION (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 664 664 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 743 743 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1323 1323 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1386 1386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 706 706 T -> A (IN REF. 5).
FT CONFLICT 1700 1700 T -> S (IN REF. 3).
FT CONFLICT 1709 1709 R -> H (IN REF. 3).
FT CONFLICT 1731 1731 S -> R (IN REF. 3).
SQ SEQUENCE 1737 AA; 192161 MW; 67FA7BF427A3DDFA CRC64;
Query Match 72.9%; Score 43; DB 1; Length 1737;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 NEEYEYE 10
DB 1409 NEDYEYD 1417
RESULT 3
ERFL TETH STANDARD; PRT; 435 AA.
AC Q3U8U5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eukaryotic peptide chain release factor subunit 1 (ERF1) (Eukaryotic
DE release factor 1).

GN ERFL.
 OS Tetrahymena thermophila.
 CC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
 CC Tetrahymenina; Tetrahymena.
 CC NCBI_TaxID=5911;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=9402978; PubMed=10471834;
 CC Karamyshev A.L., Ito K., Nakamura Y.;
 CC "Polypeptide release factor eRF1 from Tetrahymena thermophila: cDNA
 CC cloning, purification and complex formation with yeast eRF3.";
 CC FEBS Lett. 457:483-488 (1999).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=2113188; PubMed=11160924;
 CC Inagaki Y., Doolittle W.F.;
 CC "Class I release factors in ciliates variant genetic codes.";
 CC Nucleic Acids Res. 29:921-927 (2001).
 CC -1- FUNCTION: Directs the termination of nascent peptide synthesis
 CC (translation) in response to the termination codon UGA.
 CC In T.thermophila UAA and UAG codes for glutamine.
 CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, ONE OF WHICH BINDS GTP.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC RELEASE FACTOR 1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB026195; BAA85336.1; -.
 CC EMBL; AF298833; AAK07831.1; -.
 CC HSP; P46055; I079.
 CC InterPro; IPR004403; eRF1.
 CC InterPro; IPR005140; eRF1_1.
 CC InterPro; IPR005141; eRF1_2.
 CC InterPro; IPR005142; eRF1_3.
 CC Pfam; PF03463; eRF1_1; 1.
 CC Pfam; PF03464; eRF1_2; 1.
 CC Pfam; PF03465; eRF1_3; 1.
 CC TIGRFAMs; TIGR00108; eRF; 1.
 CC Protein biosynthesis.
 CC SQ SEQUENCE 435 AA; 49558 MW; 89013ED4C8612646 CRC64;
 CC -----
 CC Query Match 71.28; Score 42; DB 1; Length 435;
 CC Best Local Similarity 70.0%; Pred. No. 12;
 CC Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC QY 1 PNEEYEEYE 10
 CC |||||
 CC 422 PNEEYNEEE 431
 CC Db
 CC -----
 CC RESULT 4
 CC YG56_ARCFU
 CC ID YG56_ARCFU STANDARD; PRT; 144 AA.
 CC AC O28617;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Hypothetical protein AF1656 precursor.
 CC GN AF1656
 CC OS Archaeoglobus fulgidus.
 CC OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 CC OC Archaeoglobaceae; Archaeoglobus.
 CC OX NCBI_TaxID=2234;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 CC MEDLINE=98049343; PubMed=9389475;
 CC RX

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervagge A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Howman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370 (1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000989; AAB89614.1; -.
 CC F1R; G69456; G69456.
 CC TIGR; AF1656; -.
 CC KW Hypothetical protein; Signal; Complete proteome.
 CC FT SIGNAL 1 23 POTENTIAL.
 CC CHAIN 24 144 HYPOTHETICAL PROTEIN AF1656.
 CC SQ SEQUENCE 144 AA; 16512 MW; 81DBE3F44E477D0D CRC64;
 CC -----
 CC Query Match 66.1%; Score 39; DB 1; Length 144;
 CC Best Local Similarity 75.0%; Pred. No. 12;
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 1 PNEEYEEYE 8
 CC |||||
 CC 64 PNEEYSYD 71
 CC Db
 CC -----
 CC RESULT 5
 CC MODA_BPT4
 CC ID MODA_BPT4 STANDARD; PRT; 207 AA.
 CC AC P39423; P39422; Q38604;
 CC DT 01-FEB-1995 (Rel. 31, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE NAD+-protein ADP-ribosyltransferase modA (EC 2.4.2.-) (RNA polymerase
 CC ADP-ribosylase modA).
 CC GN MODA.
 CC OS Bacteriophage T4.
 CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 CC OC T4-like viruses.
 CC OX NCBI_TaxID=10665;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=97336891; PubMed=9193638;
 CC Wilkens K., Tiemann B., Bazan P., Rueger W.;
 RT "ADP-ribosylation and early transcription regulation by bacteriophage
 RT T4.";
 RL Adv. Exp. Med. Biol. 419:71-92 (1997).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=90255970; PubMed=1692800;
 CC Frazier M.W., Mosig G.;
 RT "The bacteriophage T4 gene mth whose product inhibits late T4 gene
 RT expression in an Escherichia coli rpoH (sigma 32) mutant.";
 RL Gene 88:7-14 (1990).
 CC [3]
 CC RP SEQUENCE FROM N.A.
 CC RP Kutter E., Arisaka F., Kunisawa T., Teugita A., Mosig G.,
 RA Mesyanzhinov V., Ruger W., Sidham T., Thomas E.;
 RT "Bacteriophage T4 genome analysis.";

```
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20097786; PubMed=10634320;
RA Tiemann B., Depping R., Rueger W.;
RT "Overexpression, purification, and partial characterization of
RT ADP-ribosyltransferases modA and modB of bacteriophage T4."
RL Gene Expr. 8:187-196(1999).
CC -!- FUNCTION: Regulates transcription by ADP-ribosylation of host RNA
CC polymerase subunit alpha.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98695; CAA67254.1; --
DR EMBL; M30001; AAB07801.1; --
DR EMBL; AF158101; AAD42602.1; --
DR PIR; T10142; T10142.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 207 AA; 24242 MW; 8C00DEC8503FA7EE CRC64;
-----
Query Match 66.1%; Score 39; DB 1; Length 207;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
-----
QY 1 PNEYEYEEYE 10
DB 143 PNEYFKTYE 152
-----
RESULT 6
UN84 CAEEL
ID UN84 CAEEL STANDARD; PRT; 1111 AA.
AC Q20745; Q9U475; Q9U476;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nuclear migration and anchoring protein unc-84 (Uncoordinated protein
DE 84).
DE UN84 OR F54B11.3.
GN Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
SEQUENCE FROM N.A. (ISOFORMS A AND B), FUNCTION, CHARACTERIZATION, AND
RP MUTANTS N321; N323; N371; N399; E1410; E1411 AND SA61.
RP STRAIN=Bristol N2; TISSUE=Embryo;
RC MEDLINE=99307080; PubMed=10375507;
RX Malone C.J., Fixsen W.D., Horvitz H.R., Han M.;
RA "UNC-84 localizes to the nuclear envelope and is required for nuclear
RT migration and anchoring during C. elegans development."
RT Development 126:3171-3181(1999).
[2]
SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Swinburne J.;
RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[3]
REVISIONS, AND ALTERNATIVE SPLICING.
RP Durbin R.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
[4]
INTERACTION WITH UNC-83.
RX MEDLINE=21614706; PubMed=11748140;
RA Sqarr D.A., Hermann G.J., Malone C.J., Fixsen W., Priess J.R.,
RA Horvitz H.R., Han M.;
RT "unc-83 encodes a novel component of the nuclear envelope and is
RT essential for proper nuclear migration."
RT Development 128:5039-5050(2001).
[5]
INTERACTION WITH ANC-1.
RX MEDLINE=22264052; PubMed=12169658;
RA Starr D.A., Han M.;
RT "Role of ANC-1 in tethering nuclei to the actin cytoskeleton."
RL Science 298:406-409(2002).
[6]
SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX MEDLINE=21904796; PubMed=11907270;
RA Lee K.K., Starr D., Cohen M., Liu J., Han M., Wilson K.L.,
RA Gruenbaum Y.;
RT "Lamin-dependent localization of UNC-84, a protein required for
RT nuclear migration in Caenorhabditis elegans."
RL Mol. Biol. Cell 13:892-901(2002).
CC -!- FUNCTION: Involved in nuclear migration and anchoring. Not
CC required for centrosome attachment to the nucleus. Probably
CC anchors the structural protein anc-1 to the nucleus, creating a
CC bridge across the nuclear envelope between the cytoskeleton and
CC the nucleus. Probably involved in nuclear migration via its
CC interaction with unc-83. Essential for the recruitment of anc-1
CC and unc-83 to the nuclear envelope.
CC -!- SUBUNIT: Interacts with unc-83 via its unc-84 domain. Interacts
CC indirectly with anc-1. Probably interacts with lamin via its N-
CC terminal domain.
CC -!- SUBCELLULAR LOCATION: Nuclear transmembrane protein; probably
CC associated with the inner nuclear membrane of the nucleus.
CC Associated with nuclei during interphase, prophase, prometaphase,
CC metaphase and early anaphase. Released from nuclear membrane in
CC the same time that the nuclear envelope disassembles, during late
CC anaphase, and begins to reaccumulate in early telophase.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=Q20745-1; Sequence=Displayed;
CC Name=b;
CC IsoId=Q20745-2; Sequence=VSP_007081, VSP_007082;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- DEVELOPMENTAL STAGE: Expressed in all cells of embryos from the
CC 26-cell stage. Then, it is ubiquitously expressed throughout the
CC development.
CC -!- DOMAIN: The UNC84 domain probably plays a role in the nuclear
CC anchoring and/or migration. Required for the localization at the
CC nuclear membrane of unc-83 and anc-1.
CC -!- SIMILARITY: Contains 1 Unc84 (SUN) domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF200706; AAF15883.1; --
EMBL; AF200707; AAF15884.1; --
EMBL; Z70208; CAA34142.2; --
EMBL; Z70208; CAC42306.1; --
DR WormPep; F54B11.3a; CE28236.
DR WormPep; F54B11.3b; CE27761.
KW Cytoskeleton; Transmembrane; Nuclear protein; Developmental protein;
KW Alternative splicing.
FT DOMAIN 1 509 NUCLEAR (POTENTIAL).
FT TRANSMEM 510 530 POTENTIAL.
FT DOMAIN 531 1111 PERINUCLEAR SPACE (POTENTIAL).
FT DOMAIN 994 1111 UNC84.
FT VARSPPLIC 877 879 LRA -> VTN (in isoform b).
FT FTId=VSP_007081.
FT Missing (in isoform b).
FT FTId=VSP_007082.
FT P->S: IN E1411; DEFECTS IN NUCLEAR
FT MUTAGEN 91 91 MIGRATION AND ANCHORING.
FT
```



```

FT MUTAGEN 932 932 D->N: IN N323; DEFECTS IN NUCLEAR
FT MUTAGEN 984 984 R->K: IN N371; DEFECTS IN NUCLEAR
FT MUTAGEN 988 988 S->F: IN SA61; DEFECTS IN NUCLEAR
FT MUTAGEN 1002 1002 G->D: IN N321 AND N399; DEFECTS IN
SQ SEQUENCE 1111 AA; 125860 MW; 6A07438E2BDC8BA6 CRC64;

Query Match 66.1%; Score 39; DB 1; Length 1111;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
Db 370 PNEYEYEEYE 379

RESULT 7
VIT2_FUNHE STANDARD; PRT; 1687 AA.
AC Q98893;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE Vitellogenin II precursor (VTG II) [Contains: Lipovitellin 1 (LV1);
DE Phosvitin (PV); Lipovitellin 2 (LV2); YP 69].
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-35.
RC TISSUE=Liver;
RA Lafleur G.J. Jr., Byrne B.M., Haux C., Greenberg R.M., Wallace R.A.;
RT "Liver-derived cDNAs: vitellogenins and vitelline envelope protein
RT precursors (choriogenins).";
RL Int. Symp. Reprod. Physiol. Fish 5:336-338(1995).
CC -!- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE
CC SOURCES OF NUTRIENTS DURING EARLY DEVELOPMENT OF OVIPAROUS
CC ORGANISMS.
CC -!- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD
CC AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENOCYTOSIS INTO GROWING
CC OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE
CC RESPECTIVE YOLK COMPONENTS LIPOVITELLINS AND PHOSVITIN.
CC -!- INDUCTION: BY STEROIDS (ESTROGEN). EXPRESSION OF VTG II IS LOWER
CC THAN THAT OF VTG I.
CC -!- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST
CC HIGHLY PHOSPHORYLATED (10%) PROTEINS IN NATURE (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U70826; AAB1752.1; -.
CC PIR; T43144; T43144.
CC InterPro; IPR001747; Lipid transprt_N.
CC DR InterPro; IPR001846; VWF_D.
CC DR Pfam; PF01347; Vitellogenin_N; 1.
CC DR Pfam; PF00094; vwd; 1.
CC DR SMART; SM00638; LPD_N; 1.
CC DR SMART; SM00216; VWD; 1.
CC KW Glycoprotein; Phosphorylation; Storage protein; Signal.
FT SIGNAL 1 15 POTENTIAL
FT CHAIN 16 1687 ? VITELLOGENIN II.
FT PIR; A81300; A81300. LIPOVITELLIN 1.

```

```

FT CHAIN ? ? ?
FT CHAIN 1059 1062 PHOSVITIN.
FT DOMAIN 1088 1169 LIPOVITELLIN 2.
FT CARBOHYD 941 941 SER-RICH.
FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1004 1004 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1019 1019 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1083 1083 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1142 1142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1179 1179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1257 1257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1292 1292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1342 1342 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1361 1361 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1390 1390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1577 1577 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1655 1655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1687 AA; 186005 MW; 4965B9B9DFB4928F CRC64;

Query Match 66.1%; Score 39; DB 1; Length 1687;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
Db 26 PGKTYEYKVE 35

RESULT 8
DAPF_CAMJE STANDARD; PRT; 249 AA.
AC Q3PMD8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Diaminopimelate epimerase (EC 5.1.1.7) (DAP epimerase).
GN DAPF OR CJ1531.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- CATALYTIC ACTIVITY: L-L-2,6-diaminoheptanedioate = meso-
CC diaminopimelate.
CC -!- PATHWAY: Biosynthesis of lysine from aspartate semialdehyde; sixth
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the diaminopimelate epimerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; ALI39078; CAB73947.1; -.
CC PIR; A81300; A81300.

```

DR HAMAP; MF_00197; -, 1.
 DR InterPro; IPR001653; DAP epimerase.
 DR Pfam; PF01678; DAP epimerase; 2.
 DR PROSITE; PS01326; DAP EPIMERASE; 1.
 KW Isomerase; Lysine biosynthesis; Complete proteome.
 FT ACT_SITE 69 BY SIMILARITY.
 FT ACT_SITE 192 BY SIMILARITY.
 SQ SEQUENCE 249 AA; 28410 MW; C6CA06E673F17348 CRC64;

Query Match 64.4%; Score 39; DB 1; Length 249;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYVEYE 10
 |.:|.:|.:|
 DB 48 PHEKYDFEWE 57

RESULT 9

LAFT VIBPA
 ID LAFT VIBPA STANDARD; PRT; 285 AA.
 AC Q03477;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Chemotaxis laft protein.
 GN LAFT OR VPA1556.
 OS Vibrio parahaemolyticus.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 CC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=BB22;
 RX MEDLINE=93273702; PubMed=8501040;
 RA McCarter L.L., Wright M.E.;

RT "Identification of genes encoding components of the swarmer cell
 flagellar motor and propeller and a sigma factor controlling
 differentiation of Vibrio parahaemolyticus.";
 RL J. Bacteriol. 175:3361-3371(1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=RIND 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).

CC -1- FUNCTION: REQUIRED FOR ROTATION OF THE FLAGELLAR MOTOR. PROBABLE
 CC TRANSMEMBRANE PROTON CHANNEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- INDUCTION: Under conditions in which the polar flagellum is not
 CC functional.

CC -1- SIMILARITY: BELONGS TO THE MOTA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U20541; AAA62354.1; -;
 CC EMBL; U52957; AAB07357.1; -;
 CC EMBL; AP005089; BAC62899.1; -;
 CC PIR; E40590; E40590.
 CC InterPro; IPR000540; Flag Mota.
 CC SuperPro; IPR002898; Mota_ExtB.
 CC Pfam; PF01618; Mota_ExtB; 1.
 CC PROSITE; PS01307; MOTA; 1.

KW Chemotaxis; Flagella; Transmembrane; Inner membrane;
 KW Flagellar rotation; Hydrogen ion transport; Complete proteome.
 FT TRANSMEM 4 23 POTENTIAL.
 FT TRANSMEM 34 51 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 201 222 POTENTIAL.
 FT DOMAIN 223 285 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 285 AA; 31324 MW; 56C6210F8A8DB652 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 285;
 Best Local Similarity 70.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PNEEYVEYE 10
 |.:|.:|.:|
 DB 68 PTEEYVEYWE 77

RESULT 10

BXCN CLOBO
 ID BXCN CLOBO STANDARD; PRT; 1196 AA.

AC P46031;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Botulinum neurotoxin type CI, nontoxic component.

OS Clostridium botulinum.
 CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.

OX NCBI_TaxID=1491;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Type C Stockholm / C-ST;

RX MEDLINE=92231894; PubMed=1567404;

RA Tezuka K., Kimura K., Fujii N., Yokosawa N., Oguma K.;

RT "The complete nucleotide sequence of the gene coding for the
 nontoxic-nonhemagglutinin component of Clostridium botulinum type C
 progenitor toxin.";

RL Biochem. Biophys. Res. Commun. 183:1273-1279(1992).

CC -1- FUNCTION: THE NONTXIC COMPONENT IS NECESSARY TO MAINTAIN

CC TOXICITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X62389; CAA44262.1; -;
 CC InterPro; IPR000395; Bontoxilysin.
 CC Pfam; PF01742; Peptidase M27; 1.
 CC PRINTS; PR00760; BONTOXILYSIN.
 CC ProDom; PD001963; Bontoxilysin; 1.
 CC Neurotoxin.

CC SEQUENCE 1196 AA; 138740 MW; 4BD5956274D7F9C3 CRC64;

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

DR

16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein MJ0070.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Xlenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073 (1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67464; AAB98053.1; -
DR PIR; F64308; F64308.
DR TIGR; MJ0070; -
DR DOMAIN 46 58 GLU-RICH
FT SEQUENCE 82 AA; 9607 MW; 29823047B2C550B CRC64;
SQ
Query Match 62.7%; Score 37; DB 1; Length 82;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 PNEEYEEYE 10
DB 45 PEEEFEDYE 54
RESULT 12
Y665_CHLMU
ID Y665_CHLMU STANDARD; PRT; 289 AA.
AC Q9PK08;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein Y665.
DN TC0665.
GN Chlamydia muridarum.
OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
White O., Hickey E.K., Peterson J., Utterback T., Berry K., Dodson R.,
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Salzborg S.L.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0160 (MTG1) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002334; AAF39487.1; ALT_INIT.
DR TIGR; TC0665; -
DR InterPro; IPR003226; Metdp_prot_hydro.
DR Pfam; PF03690; UPF0160; 1.
DR ProDom; PD034736; Metdp_prot_hydro; 1.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 289 AA; 33067 MW; 524718A28D229762 CRC64;
Query Match 62.7%; Score 37; DB 1; Length 289;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 PNEEYEEY 7
DB 93 PHEEYEEY 99
RESULT 13
HEM1_THEVO
ID HEM1_THEVO STANDARD; PRT; 409 AA.
AC Q97B68;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamy1-tRNA reductase (EC 1.2.1.-) (GLUTR).
GN HEMA OR TV0590 OR TVG0579613.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatiales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GSS1 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
sequence of Thermoplasma volcanium.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262 (2000).
CC -1- CATALYTIC ACTIVITY: Glutamy1-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -1- SIMILARITY: Belongs to the glutamy1-tRNA reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000992; BAB59732.1; -
DR HAMAP; MF 00087; -; 1.
DR InterPro; IPR000343; GlutR.
DR Pfam; PF00745; GlutR_dimer; 1.
DR Pfam; PF05201; GlutR_N; 1.
DR Pfam; PF05200; GlutR_NAD_bind; 1.
DR TIGRFAMs; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; FALSE NEG.
KW Porphyrin biosynthesis; Oxidoreductase; NADP: Complete proteome.
FT ACT_SITE 49 49 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 79 79 BASE (BY SIMILARITY).
SQ SEQUENCE 409 AA; 46691 MW; 8CF23510B233D830 CRC64;
CC

Query Match 62.7%; Score 37; DB 1; Length 409;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EYEYEEY 9
|:|||||
Db 199 EYEYEEY 205

RESULT 14
FLO3_CABEL STANDARD; PRT; 424 AA.
AC O45166;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Folate-like transporter 3.
GN F37B4.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Henkhaus J., Wohlmann P., O'Brien D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE SLC19A FAMILY OF TRANSPORTERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF047657; AAK18943.1; -
DR PIR; A88979; A88979.
DR WormPep; F37B4.7; CBI7796.
DR InterPro; IPR002666; Folate carrier.
DR Pfam; PF01770; Folate carrier; 1.
DR TIGRFAMs; TIGR00806; ffc; 1.
KW Folate-binding; Transport; Transmembrane; Glycoprotein;
KW Multigene family.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 392 412 POTENTIAL.
SQ SEQUENCE 424 AA; 48964 MW; 87386CE6C5C54878 CRC64;

Query Match 62.7%; Score 37; DB 1; Length 424;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|:|||||
Db 21 PTEPYMYEQ 30

RESULT 15
FLO2_CABEL STANDARD; PRT; 535 AA.
ID FLO2_CABEL
AC Q22931;
DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Folate-like transporter 2.
GN CS0E3.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisel C., Bradshaw H.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE SLC19A FAMILY OF TRANSPORTERS.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U64848; AAB04886.1; -
DR PIR; T29677; T29677.
DR WormPep; CS0E3.11; CR08904.
DR InterPro; IPR002666; Folate carrier.
DR Pfam; PF01770; Folate carrier; 1.
DR TIGRFAMs; TIGR00806; ffc; 1.
KW Folate-binding; Transport; Transmembrane; Glycoprotein;
KW Multigene family.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 77 97 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 211 231 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
FT TRANSMEM 339 359 POTENTIAL.
SQ SEQUENCE 535 AA; 62279 MW; 408EB1CDB3D29C6B CRC64;

Query Match 62.7%; Score 37; DB 1; Length 535;
Best Local Similarity 60.0%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|:|||||
Db 21 PTEPYMYEQ 30

Search completed: August 5, 2003, 07:38:54
Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:36:49 ; Search time 33 Seconds
(without alignments)
78.198 Million cell updates/sec

Title: US-09-883-727A-125
Perfect score: 59
Sequence: 1 PNEYEYEEYE 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 23:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_ivirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	78.0	247	10 Q9AYJ6	Q9ayj6 oryza sativ
2	44	74.6	391	5 Q26258	Q26258 chironomus
3	44	74.6	511	5 Q00625	Q00625 chironomus
4	44	74.6	561	5 Q95P96	Q95P96 patinopecte
5	44	74.6	1238	13 Q8JIF8	Q8jif8 acanthogobi
6	44	74.6	7170	5 Q8IIL30	Q8il30 plasmodium
7	43	72.9	163	5 Q9NDN8	Q9ndn8 armadillidi
8	43	72.9	749	5 Q23804	Q23804 chironomus
9	43	72.9	1088	12 Q9QNB3	Q9qnb3 human rotav
10	43	72.9	1088	12 Q85036	Q85036 porcine rot
11	43	72.9	2763	5 Q8I148	Q8i148 plasmodium
12	42	71.2	149	16 Q31728	Q31728 bacillus su
13	42	71.2	267	5 Q26259	Q26259 chironomus
14	42	71.2	387	12 Q83389	Q83389 unidentified
15	42	71.2	451	5 Q18377	Q18377 caenorhabdi
16	41	69.5	218	16 Q9X8F5	Q9x8f5 streptomyce

ALIGNMENTS

RESULT 1

ID	Q9AYJ6	PRELIMINARY;	PRT;	247 AA.
AC	Q9AYJ6;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	Putative DNA-binding protein (Putative homeobox protein HAT22).			
GN	OSJNBA0071K19.2 OR OSJNBA0096B22.15.			
OS	Oryza sativa (Rice), and			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzaceae; Oryza.			
OX	NCBI_TaxID=4530, 39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Wing R.A., Frisch D., Presting G., Wood T., Yu Y., Soderlund C.,			
RA	Kim H., Rambo T., Henry D., Simmons J., Wilson R., Johnson D.,			
RA	Bradshaw H., Du H.;			
RT	"Rice Genomic Sequence."			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,			
RA	Saski C., Henry D., Oates R., Simmons J.;			
RT	"Rice Genomic Sequence."			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
DR	EMBL; AC069324; AAK00416.1; -;			
DR	EMBL; AC099400; AAL91609.1; -;			
DR	HSSP; P01366; 1AKH.			
DR	Gramene; Q9AYJ6; -;			
DR	InterPro; IPR001356; Homeobox.			
DR	InterPro; IPR003106; Leu_zip_homeo.			
DR	Pfam; PF02183; HALZ; 1.			
DR	Pfam; PF00046; homeobox; 1.			
DR	ProDom; PD000010; Homeobox; 1.			
DR	SMART; SM00340; HALZ; 1.			

Q8r974 thermoanaer
O05579 mycobacteri
Q8vka2 mycobacteri
Q8ilq0 plasmodium
Q8id46 plasmodium
Q8f7r5 leptospira
Q92jm8 rickettsia
Q9dhi7 yaba-like d
Q97u45 sulfolobus
Q9vvg7 drosophila
Q8i2m8 plasmodium
Q9kb49 bacillus ha
Q95x10 caenorhabdi
Q8iu34 crassostrea
Q8ihp8 plasmodium
Q8x17 pyrobaculum
Q8y9p7 listeria mo
Q22513 caenorhabdi
Q9lix9 oryza sativ
Q9u475 caenorhabdi
Q98701 elm mottle
Q20745 caenorhabdi
Q8uw61 oryzias lat
Q62263 caenorhabdi
Q8zmh6 salmoneila
Q8er44 oceanobacil
Q8rael thermoanaer
Q8sdu0 bacterioph
Q93lx4 staphylococ

17 41 69.5 319 16 Q8R974
18 41 69.5 358 16 O05579
19 41 69.5 391 16 Q8VKA2
20 41 69.5 484 5 Q8ILQ0
21 41 69.5 3482 5 Q8ID46
22 40 67.8 156 16 Q8F7R5
23 40 67.8 222 16 Q92JM8
24 40 67.8 223 12 Q9DHI7
25 40 67.8 259 17 Q97U45
26 40 67.8 348 5 Q9VMG7
27 40 67.8 405 5 Q8I2M8
28 40 67.8 500 16 Q9KB49
29 40 67.8 1142 5 Q95XL0
30 40 67.8 1583 5 Q8IU34
31 40 67.8 2755 5 Q8IHP8
32 39 66.1 59 17 Q8X17
33 39 66.1 125 16 Q8Y9P7
34 39 66.1 347 5 Q22513
35 39 66.1 641 10 Q9LIX9
36 39 66.1 879 5 Q9U475
37 39 66.1 1055 12 Q98701
38 39 66.1 1111 5 Q20745
39 39 66.1 1725 13 Q8UW61
40 39 66.1 3396 5 Q62263
41 38 64.4 108 16 Q8ZMH6
42 38 64.4 144 16 Q8ER44
43 38 64.4 158 16 Q8RAE1
44 38 64.4 193 9 Q8SDU0
45 38 64.4 193 16 Q93LX4

DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 247 AA; 2718 AA; 2778 MW; D13AE4C477E74AOA CRC64;

Query Match 77.8%; Score 46; DB 10; Length 247;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYEYEYE 10
 Db 66 EYEYEYE 73

RESULT 2

ID Q26258 PRELIMINARY; PRT; 391 AA.
 AC Q26258;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE BR2.2 protein (Fragment).
 GN BR2.2.
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomoidea;
 OC Chironomidae; Chironominae; Chironomus.
 OX NCBI_TaxID=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92389373; PubMed=1518088;
 RA Paulsson G., Bernholm K., Wieslander L.;
 RT "Conserved and variable repeat structures in the Balbiani ring gene
 family in Chironomus tentans.";
 RL J. Mol. Evol. 35:205-216(1992).
 DR EMBL; S45043; AAB23117.1; -.
 FT NON TER 391 391
 SQ SEQUENCE 391 AA; 43405 MW; DCEC9ED37FE8B756 CRC64;

Query Match 74.6%; Score 44; DB 5; Length 391;
 Best Local Similarity 77.8%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYEYEYE 10
 Db 102 DEYEYEYE 110

RESULT 3

ID Q00625 PRELIMINARY; PRT; 511 AA.
 AC Q00625;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE BR1 protein precursor (Fragment).
 GN BR1.
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomoidea;
 OC Chironomidae; Chironominae; Chironomus.
 OX NCBI_TaxID=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92277646; PubMed=1593624;
 RA Paulsson G., Hoog C., Bernholm K., Wieslander L.;
 RT "Balbiani ring 1 gene in Chironomus tentans. Sequence organization and
 dynamics of a coding miniatellite.";
 RL J. Mol. Biol. 225:349-361(1992).
 DR EMBL; X64322; CAA45607.1; -.
 KW Signal.
 FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 >511 POTENTIAL.
 FT NON TER 511 511
 SQ SEQUENCE 511 AA; 56077 MW; 8ABE75847FD2ECB0 CRC64;

Query Match 74.8%; Score 44; DB 5; Length 511;
 Best Local Similarity 77.8%; Pred. No. 26;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NEEYEYEYE 10
 Db 111 DEYEYEYE 119

RESULT 4

ID Q95P96 PRELIMINARY; PRT; 561 AA.
 AC Q95P96;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Vitellogenin (Fragment).
 GN VTG.
 OS Patinopecten yessoensis (Ezo giant scallop) (Yesso scallop).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
 OC Pectinoidea; Pectinidae; Mizuhopecten.
 OX NCBI_TaxID=6573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Osada M.;
 RT "Patinopecten yessoensis mRNA for vitellogenin, partial cds.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB055960; BAB63260.1; -.
 DR InterPro; IPR001747; Lipid_transprt_N.
 DR Pfam; PF01347; Vitellogenin_N; 1.
 DR SMART; SM00638; LPD_N; 1.
 FT NON TER 561 561
 SQ SEQUENCE 561 AA; 63987 MW; C939FBECAB7907B1 CRC64;

Query Match 74.6%; Score 44; DB 5; Length 561;
 Best Local Similarity 70.0%; Pred. No. 28;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYEYEYE 10
 Db 26 PNKEYWYDYE 35

RESULT 5

ID Q8JIF8 PRELIMINARY; PRT; 1238 AA.
 AC Q8JIF8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Vitellogenin.
 GN VG-320.
 OS Acanthogobius flavimanus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;
 OC Gobiidae; Acanthogobius.
 OX NCBI_TaxID=86203;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Okubo N., Mochida K., Adachi S., Hara A., Matsubara T.;
 RT "Deduced primary structures of two forms of vitellogenin in Japanese
 common goby.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB088474; BAC06191.1; -.
 DR InterPro; IPR001747; Lipid_transprt_N.
 DR Pfam; PF01347; Vitellogenin_N; 1.
 DR SMART; SM00638; LPD_N; 1.

SQ SEQUENCE 1238 AA; 140519 MW; CA1275F6A8376803 CRC64;
Query Match 74.6%; Score 44; DB 13; Length 1238;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PNEEYEEYE 10
||: |||: ||
Db 27 PNKIYEYQE 36

RESULT 6
Q8IL30 PRELIMINARY; PRT; 7170 AA.
ID Q8IL30
AC Q8IL30;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF14_0419.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7.
RX MEDLINE=22255705; PubMed=12369864;
RA Gardner M.J., Hall N., Fung E., White O., Bertman M., Hyman R.W.,
RA Carlton J.M., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014823; AAN37032.1; -.
KW Hypothetical protein.
SQ SEQUENCE 7170 AA; 857927 MW; 4F15304556D54528 CRC64;

Query Match 74.6%; Score 44; DB 5; Length 7170;
Best Local Similarity 77.8%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 NEEYEEYE 10
::: |||: |||
Db 5759 SDEYEEYE 5767

RESULT 7
Q9NDN8 PRELIMINARY; PRT; 163 AA.
ID Q9NDN8
AC Q9NDN8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Vitellogenin (Fragment).
GN VTG.
OS Armadillidium vulgare (Woodlice) (Pillbugs).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Petacarida; Isopoda; Oniscidea; Armadillidiidae;
OC Armadillidium.
OX NCBI_TaxID=13347;
RN [1]
RP SEQUENCE FROM N.A.
RA Okuno A., Katayama H., Nagasawa H.;
RT "Partial characterization of vitellin and localization of vitellogenin
RT production in the terrestrial isopod, Armadillidium vulgare."
RL Comp. Biochem. Physiol. 126:397-407(2000).
DR EMBL; AB037247; BAB03280.1; -.

DR InterPro; IPR001747; Lipid transport_N.
DR Pfam; PF01347; Vitellogenin_N; 1.
FT NON TER 163 163
SQ SEQUENCE 163 AA; 18261 MW; DA5199171E6CE0EF CRC64;
Query Match 72.9%; Score 43; DB 5; Length 163;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PNEEYEEYE 10
|: |||: ||
Db 19 PGMEYEDYE 28

RESULT 8
Q23804 PRELIMINARY; PRT; 749 AA.
ID Q23804
AC Q23804;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Spid precursor (Fragment).
GN BR2.1.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;
OC Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92262483; PubMed=1584794;
RA Wieslander L., Paulsson G.;
RT "Sequence organization of the Balbiani ring 2.1 gene in Chironomus
RT tentans."
RL Proc. Natl. Acad. Sci. U.S.A. 89:4578-4582(1992).
DR EMBL; M89909; AAA28264.1; -.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT NON TER 749 749
SQ SEQUENCE 749 AA; 83564 MW; 606367E2800C45BD CRC64;

Query Match 72.9%; Score 43; DB 5; Length 749;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 EEEYEEYE 10
::: |||: |||
Db 122 DEYEEYE 129

RESULT 9
Q9QNB3 PRELIMINARY; PRT; 1088 AA.
ID Q9QNB3
AC Q9QNB3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE VP1.
GN VP1.
OS Human rotavirus (strain KU).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KU;
RA Taniguchi K.;
RT "Rotavirus."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022765; BAA84962.1; -.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007097; RNA_pol_reo.
RL PROSITE; P850523; RDRF_REO; 1.
DR PROSITE; P850521; RDRF_VIRAL; 1.

SQ SEQUENCE 1088 AA; 125016 MW; 214CA10B5825508 CRC64;

Query Match 72.9%; Score 43; DB 12; Length 1088;

Best Local Similarity 77.8%; Pred. No. 81;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NEEYEYE 10

Db 186 NEKYRYE 194

RESULT 10

ID Q85036 PRELIMINARY; PRT; 1088 AA.

AC Q85036;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE VP1 protein.

GN 1.

OS Porcine rotavirus.

OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.

OX NCBI_TaxID=10913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95140933; PubMed=7839009;

RA Almanza L., Arias C.F., Lopez S.;

RT "Aminoacid sequence of the Porcine rotavirus YM VP1 protein.";

RL Res. Virol. 145:313-317(1994).

DR EMBL; X76486; CAA54024.1; -.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR007097; RNA_pol_reo.

DR PROSITE; PS50523; RDRP_REO.1.

DR PROSITE; PS50523; RDRP_VIRAL; 1.

SQ SEQUENCE 1088 AA; 124792 MW; 73C5723B9EA74357 CRC64;

Query Match

Best Local Similarity 72.9%; Score 43; DB 12; Length 1088;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NEEYEYE 10

Db 186 NEKYRYE 194

RESULT 11

ID Q81148 PRELIMINARY; PRT; 2763 AA.

AC Q81148;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Hypothetical protein.

GN PF11_0326.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3D7;

RX MEDLINE=22255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,

RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,

RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,

RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,

RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,

RA Fraser C.M., Barrell B.,

RT "Genome sequence of the human malaria parasite Plasmodium

RT falciparum.";

RL Nature 419:498-511(2002).

DR EMBL; AE014840; AAN35910.1; -.

KW Hypothetical protein.

SQ SEQUENCE 2763 AA; 330300 MW; 23AFOD9286193774 CRC64;

Query Match 72.9%; Score 43; DB 5; Length 2763;

Best Local Similarity 87.5%; Pred. No. 21e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EEEYEYE 10

Db 691 QEEYEYE 698

RESULT 12

ID Q31728 PRELIMINARY; PRT; 149 AA.

AC Q31728;

DT 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE YLMP protein.

GN YLMP.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Berto M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Bruchet S., Bruchet C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denist F., Devine K.M., Dusterhoft A., Enrich S.D., Emmerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrar E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi E., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the gram-positive bacterium Bacillus

RT subtilis.";

RL Nature 390:249-256(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z99112; CAB13413.1; -.

KW Complete proteome.

SQ SEQUENCE 149 AA; 17155 MW; A78E3576FD19A0EA CRC64;

Query Match 71.2%; Score 42; DB 16; Length 149;

Best Local Similarity 87.5%; Pred. No. 16;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEEYEY 9
:|||||
Db 13 DEEY 20

RESULT 13

Q26259 PRELIMINARY; PRT; 267 AA.
ID Q26259
AC Q26259;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BR6 protein (Fragment).
GN BR6.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;
OC Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92389373; PubMed=1518088;
RA Paulsson G., Bernholm K., Wieslander L.;
RT "Conserved and variable repeat structures in the Balbiani ring gene
family in Chironomus tentans.";
RL J. Mol. Evol. 35:205-216(1992).
DR EMBL; S45044; AAB23118.1; -.
FT NON TER 267
SQ SEQUENCE 267 AA; 1974 MW; BC4D1615E5AA0E24 CRC64;

Query Match 71.2%; Score 42; DB 5; Length 267;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NEEYEY 10
:|||||
Db 104 DNEYEY 112

RESULT 14

Q83389 PRELIMINARY; PRT; 387 AA.
ID Q83389
AC Q83389;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Fusolin precursor.
OS unidentified entomopoxvirus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae.
OX NCBI_TaxID=10291;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266248; PubMed=7747415;
RA Gauthier L., Veyrunes J.C., Cousseans F., Bergoin M.;
RT "The Melonthea melonthea Entomopoxvirus (MmEPV) Fusolin Is Related
To The 50K EPV Polypeptide And To The Baculovirus 37K Glycoprotein.";
RL Virology 208:427-436(1995).
DR EMBL; X77616; CAA54706.1; -.
DR InterPro; IPR004302; Chitin binding_3.
DR Pfam; PF03067; Chitin_bind_3; 1.
KW Signal.
FT SIGNAL
SQ SEQUENCE 387 AA; 45834 MW; CDC14D9FDEC37424 CRC64;

Query Match 71.2%; Score 42; DB 12; Length 387;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NEEYEY 10
:|||||
Db 316 NYEY 324

RESULT 15

Q18377 PRELIMINARY; PRT; 451 AA.
ID Q18377
AC Q18377;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C33D9.6 protein.
GN C33D9.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z68159; CAA92287.4; -.
DR WormPep; C33D9.6; CE29588.
SQ SEQUENCE 451 AA; 52769 MW; 83792CAC8B5C09D2 CRC64;

Query Match 71.2%; Score 42; DB 5; Length 451;
Best Local Similarity 77.8%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEY 9
:|||||
Db 420 PEEY 428

Search completed: August 5, 2003, 07:39:34
Job time : 35 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: August 5, 2003, 07:34:19 ; Search time 41 Seconds
(without alignments)
38.714 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYEEYE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	10	23	AAE20072
2	56	94.9	10	23	AAE19949
3	56	94.9	10	23	AAE19957
4	56	94.9	10	23	AAE19958
5	56	94.9	10	23	AAE19959
6	56	94.9	10	23	AAE19960
7	56	94.9	10	23	AAE19961
8	56	94.9	10	23	AAE19962
9	56	94.9	10	23	AAE19963
					Complement C1s inh
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo

10	56	94.9	122	23	AAE19948	Haementaria ghilia
11	56	94.9	142	20	AAE06449	Leech haemostasin.
12	56	94.9	142	20	AAE06450	Leech haemostasin.
13	52	88.1	10	23	AAE19950	Complement C1s exo
14	52	88.1	10	23	AAE19951	Complement C1s exo
15	52	88.1	10	23	AAE19952	Complement C1s exo
16	52	88.1	10	23	AAE19964	Complement C1s exo
17	52	88.1	10	23	AAE19965	Complement C1s exo
18	52	88.1	10	23	AAE19966	Complement C1s exo
19	52	88.1	10	23	AAE19967	Complement C1s exo
20	52	88.1	10	23	AAE19968	Complement C1s exo
21	52	88.1	10	23	AAE19969	Complement C1s exo
22	52	88.1	10	23	AAE19972	Complement C1s exo
23	52	88.1	10	23	AAE19974	Complement C1s exo
24	48	81.4	10	23	AAE19953	Complement C1s exo
25	48	81.4	10	23	AAE19954	Complement C1s exo
26	48	81.4	10	23	AAE19955	Complement C1s exo
27	48	81.4	10	23	AAE19970	Complement C1s exo
28	48	81.4	10	23	AAE19971	Complement C1s exo
29	48	81.4	10	23	AAE19973	Complement C1s exo
30	44	74.6	10	23	AAE19956	Complement C1s exo
31	41	69.5	319	24	ABP97413	Thermoanaerobacter
32	41	69.5	358	23	ABU05789	M. tuberculosis an
33	41	69.5	430	22	ABE68327	Drosophila melanog
34	40	67.8	348	22	ABE69117	Listeria monocytog
35	39	66.1	125	23	ABB49064	B. burgdorferi ant
36	38	64.4	179	20	AAE19847	Staphylococcus aur
37	38	64.4	193	22	AAU37533	B. burgdorferi ant
38	38	64.4	203	20	AAE19846	Drosophila melanog
39	38	64.4	326	22	ABE68891	GTP enzyme Rho fam
40	38	64.4	816	22	AAE66505	Zebrafish transcri
41	38	64.4	1084	23	ABB76489	Drosophila melanog
42	38	64.4	1310	22	ABB60913	Drosophila melanog
43	37.5	63.6	777	22	ABB60154	Biotinylated pepti
44	37	62.7	19	21	AAE85176	Src protein kinase
45	37	62.7	19	22	AAU11804	

ALIGNMENTS

RESULT 1
AAE20072
ID AAE20072 standard; peptide; 10 AA.

XX AC AAE20072;

XX DT 18-JUN-2002 (first entry)

XX DE Complement C1s inhibitor peptide #1.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX

PT New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
PT sepsis
XX
PS Claim 4; Page 49; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejections, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s inhibitor peptide.
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 59; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PNEYEYEEYE 10
DB 1 PNEYEYEEYE 10
|||||
RESULT 2
AAE19949
ID AAE19949 standard; peptide; 10 AA.
XX
AC AAE19949;
XX
DT 18-JUN-2002 (first entry)
XX
DE Complement C1s exosite binding peptide #1.
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
OS Unidentified.
XX
FN WO200198365-A2.
XX
PD 27-DEC-2001.
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
PR 21-JUN-2000; 2000US-212998P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI West RR, Sheppard PO, Fox BA;
XX
DR WPI; 2002-241177/29.
XX
CC New complement C1s inhibitor polypeptides for treating diseases in
CC which complement activation has been shown to occur, e.g. adult
CC respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
CC sepsis
XX
PS Disclosure; Page 15; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
SQ Sequence 10 AA;
Query Match 94.9%; Score 56; DB 23; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0037;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 PNEYEYEEYE 10
DB 1 PNEYEYEEYE 10
|||||
RESULT 3
AAE19957
ID AAE19957 standard; peptide; 10 AA.
XX
AC AAE19957;
XX
DT 18-JUN-2002 (first entry)
XX
DE Complement C1s exosite binding peptide #9.
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
OS Unidentified.
XX
FN WO200198365-A2.
XX
PD 27-DEC-2001.
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
PR 21-JUN-2000; 2000US-212998P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI West RR, Sheppard PO, Fox BA;
XX
DR WPI; 2002-241177/29.
XX
CC New complement C1s inhibitor polypeptides for treating diseases in
CC which complement activation has been shown to occur, e.g. adult
CC respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
CC sepsis
XX
PS Disclosure; Page 15; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYDYDYE 10

RESULT 4
 AAE19958
 ID AAE19958 standard; peptide; 10 AA.
 AC AAE19958;
 XX 18-JUN-2002 (first entry)
 DT
 XX Complement C1s exosite binding peptide #10.
 DE
 XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute reagent; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.

XX Key Location/Qualifiers
 FH Modified-site 7 /note= "Sulphated Tyr"
 FT
 FT
 FT
 XX WO200198365-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX 18-JUN-2001; 2001WO-US19405.
 XX
 XX 21-JUN-2000; 2000US-212998P.
 PR
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX West RR, Sheppard PO, Fox BA;
 PI
 XX WPI; 2002-241177/29.
 DR
 XX
 XX New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis -
 XX
 XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute reagent, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYDYDYE 10

RESULT 5
 AAE19959
 ID AAE19959 standard; peptide; 10 AA.
 XX
 AC AAE19959;
 XX 18-JUN-2002 (first entry)
 DT
 XX Complement C1s exosite binding peptide #11.
 DE
 XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute reagent; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.

XX Key Location/Qualifiers
 FH Modified-site 9 /note= "Sulphated Tyr"
 FT
 FT
 FT
 XX WO200198365-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX 18-JUN-2001; 2001WO-US19405.
 XX
 XX 21-JUN-2000; 2000US-212998P.
 PR
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX West RR, Sheppard PO, Fox BA;
 PI
 XX WPI; 2002-241177/29.
 DR
 XX
 XX New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis -
 XX
 XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute reagent, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYDYDYE 10

RESULT 6
 AAE19960
 ID AAE19960 standard; peptide; 10 AA.
 XX
 AC AAE19960;
 XX 18-JUN-2002 (first entry)
 DT

XX DE Complement C1s exosite binding peptide #12.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute refection; rheumatoid arthritis; burn; wound healing; asthma;

XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

XX KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Modified-site 5 /note= "Sulphated Tyr"

XX FT Modified-site 7 /note= "Sulphated Tyr"

XX FT Modified-site 9 /note= "Sulphated Tyr"

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX PS New complement C1s inhibitor polypeptides for treating diseases in

XX PT which complement activation has been shown to occur, e.g. adult

XX PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or

XX PT sepsis

XX PS Disclosure; Page 15; 99pp; English.

XX CC The invention relates to a polypeptide that inhibits complement C1s. The

XX CC inhibitory peptides are useful as therapeutic agents, as preservatives

XX CC in blood samples, and in affinity purification procedures to isolate C1s.

XX CC Molecules that inhibit complement may be used for treating diseases

XX CC in which complement activation has been shown to occur, e.g., adult

XX CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial

XX CC infarct, stroke), hyperacute refection, sepsis, burns, wound healing,

XX CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

XX CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present

XX CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;

Best Local Similarity 90.0%; Pred. NO. 0.0037;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PNEEYVEYE 10

DB 1 PNEEYDYDE 10

RESULT 7

AAE19961

ID AAE19961 standard; peptide; 10 AA.

XX AC AAE19961;

XX XX

DT 18-JUN-2002 (first entry)

XX DE Complement C1s exosite binding peptide #12.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute refection; rheumatoid arthritis; burn; wound healing; asthma;

KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Modified-site 5 /note= "Sulphated Tyr"

XX FT Modified-site 9 /note= "Sulphated Tyr"

XX PN WO200198365-A2.

XX PD 27-DEC-2001.

XX PF 18-JUN-2001; 2001WO-US19405.

XX PR 21-JUN-2000; 2000US-212998P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI West RR, Sheppard PO, Fox BA;

XX DR WPI; 2002-241177/29.

XX PS New complement C1s inhibitor polypeptides for treating diseases in

XX PT which complement activation has been shown to occur, e.g. adult

XX PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or

XX PT sepsis

XX PS Disclosure; Page 15; 99pp; English.

XX CC The invention relates to a polypeptide that inhibits complement C1s. The

XX CC inhibitory peptides are useful as therapeutic agents, as preservatives

XX CC in blood samples, and in affinity purification procedures to isolate C1s.

XX CC Molecules that inhibit complement may be used for treating diseases

XX CC in which complement activation has been shown to occur, e.g., adult

XX CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial

XX CC infarct, stroke), hyperacute refection, sepsis, burns, wound healing,

XX CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

XX CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present

XX CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;

Best Local Similarity 90.0%; Pred. NO. 0.0037;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PNEEYVEYE 10

DB 1 PNEEYDYDE 10

RESULT 8

AAE19962

ID AAE19962 standard; peptide; 10 AA.

XX AC AAE19962;

XX XX

DT 18-JUN-2002 (first entry)

XX DE Complement C1s exosite binding peptide #12.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute refection; rheumatoid arthritis; burn; wound healing; asthma;

XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

XX KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX DR WPI; 2002-241177/29.

XX PT New complement C1s inhibitor polypeptides for treating diseases in

PT which complement activation has been shown to occur, e.g. adult

PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or

PT sepsis

XX PS Claim 6; Page 2; 99pp; English.

XX CC The invention relates to a polypeptide that inhibits complement C1s. The

CC inhibitory peptides are useful as therapeutic agents, as preservatives

CC in blood samples, and in affinity purification procedures to isolate C1s.

CC Molecules that inhibit complement may be used for treating diseases

CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial

CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,

CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present

CC sequence is Haementeria ghilianii BD001 protein used in the invention.

XX SQ Sequence 122 AA;

Query Match 94.9%; Score 56; DB 23; Length 122;

Best Local Similarity 90.0%; Pred. No. 0.063;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10

DB 113 PNEEYDYDYE 122

RESULT 11

AA06449 ID AA06449 standard; Protein; 142 AA.

AC AA06449;

XX DT 27-SEP-1999 (first entry)

XX DE Leech haemostasin.

XX KW Haemostasin; leech; blood clotting; coagulation; complement;

KW Factor XII; C1 esterase; inhibitor; cardiovascular disease;

KW inflammation; autoimmune disease; thrombosis; pulmonary embolus;

KW angioplasty; endarterectomy; haemodialysis; cardiopulmonary bypass;

KW organ rejection; sepsis; myocardial infarction; stroke;

KW ischaemic shock; atherosclerosis; vasculitis; rheumatoid arthritis;

KW sickle cell anaemia; angioedema; lupus arthritis; nephritis;

KW glomerulonephritis; systemic sclerosis; Behcet's syndrome;

KW cerebral lupus; Guillain-Barre disease; multiple sclerosis;

KW myasthenia gravis; pemphigus; bullous pemphigoid; burn; asthma;

KW phototoxic reaction; skin reaction; infection; thyroiditis;

KW inflammatory bowel disease; Alzheimer's disease;

KW paroxysmal nocturnal haemoglobinuria; haemolytic anaemia; therapy.

XX OS Haementeria ghilianii.

XX FH Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "leader peptide"

FT Protein 21..142

FT Protein /note= "mature protein, specifically claimed in

FT Protein Claim 5"

XX PN WO9936439-A1.

XX PD 22-JUL-1999.

XX PF 15-JAN-1999; 99WO-GB00147.

XX PR 16-JAN-1998; 98GB-0000817.

XX

PA (BIOD-) BIO-DISCOVERY LTD.

XX PI Finney S, Seale L;

XX DR WPI; 1999-430600/36.

XX DR N-PSDB; AAX59408.

XX PT New polypeptides which inhibit C1 esterase and/or factor XII

PT activation are used to treat cardiovascular, inflammatory or

PT autoimmune disorders

XX XX Claim 5; Page 5-6; 43pp; English.

XX CC The present sequence represents a haemostasin polypeptide of the

CC leech Haementeria ghilianii. The 14 kDa mature portion of the

CC polypeptide is claimed. It acts as an inhibitor of the initiation

CC complexes of blood coagulation and/or of complement. 3 Variants

CC (see also AA06450 and AA06451) of haemostasin were identified that

CC vary at positions 16 and 60 of the mature protein. The invention

CC provides vectors and host cells suitable for production of

CC haemostasin. A claimed method for the treatment or prophylaxis of

CC a condition or disorder related to C1 esterase initiation and/or

CC Factor XII activation comprises the administration to a patient of

CC an effective, inhibitory amount of a haemostasin polypeptide. The

CC polypeptide may administered by means of being expressed in the

CC cells of the patient, the cells comprising and expressing a

CC haemolysin nucleic acid sequence. Haemostasins can potentially be

CC used to inhibit the activation of coagulation in the treatment of

CC thrombotic disease such as deep venous thrombosis, pulmonary

CC embolus, and thrombosis associated with angioplasty and

CC endarterectomy. Disease may also be alleviated by the ability of the

CC haemostasins to inhibit both complement activation and the intrinsic

CC pathway of blood coagulation, such as in haemodialysis.

CC cardiopulmonary bypass, or rejection of transplanted organs or

CC tissues, or in cases of sepsis, myocardial infarction, stroke,

CC particularly in the injury caused to tissues by reperfusion after

CC an ischaemic period, atherosclerosis, shock, vasculitis,

CC rheumatoid arthritis, sickle cell anaemia or angioedema. The

CC haemostasins may also be used in conditions associated with

CC activation of complement, such as: various autoimmune diseases

CC (e.g. lupus arthritis); glomerulonephritis; nephritis; nephropathy;

CC systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillain-Barre

CC disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous

CC pemphigoid; phototoxic reactions; thermal burns; anaphylaxis;

CC asthma; skin reactions; infections; inflammatory bowel disease;

CC thyroiditis; infertility; Alzheimer's disease; paroxysmal

CC nocturnal haemoglobinuria; and haemolytic anaemia (all claimed).

XX SQ Sequence 142 AA;

Query Match 94.9%; Score 56; DB 20; Length 142;

Best Local Similarity 90.0%; Pred. No. 0.075;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10

DB 133 PNEEYDYDYE 142

RESULT 12

AA06450 ID AA06450 standard; Protein; 142 AA.

XX AC AA06450;

XX DT 27-SEP-1999 (first entry)

XX DE Leech haemostasin.

XX KW Haemostasin; leech; blood clotting; coagulation; complement;

KW Factor XII; C1 esterase; inhibitor; cardiovascular disease;

KW inflammation; autoimmune disease; thrombosis; pulmonary embolus;

KW angioplasty; endarterectomy; haemodialysis; cardiopulmonary bypass;

organ rejection; sepsis; myocardial infarction; stroke; ischaemic shock; atherosclerosis; vasculitis; rheumatoid arthritis; sickle cell anaemia; angioedema; lupus arthritis; Behcet's syndrome; glomerulonephritis; systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillan-Barre disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous pemphigoid; burn; asthma; phototoxic reaction; skin reaction; infection; thyroiditis; inflammatory bowel disease; Alzheimer's disease; paroxysmal nocturnal haemoglobinuria; haemolytic anaemia; therapy.

Haementeria ghilianii.

Key Location/Qualifiers
 Peptide 1..20
 Protein 21..142
 /note= "leader peptide"
 /note= "mature protein, specifically claimed in Claim 5"

WO9936439-A1.
 22-JUL-1999.
 15-JAN-1999; 99WO-GB00147.
 16-JAN-1998; 98GB-0000817.
 (BIOD-) BIO-DISCOVERY LTD.
 Finney S, Seale L;
 WPI; 1999-430600/36.

New polypeptides which inhibit C1 esterase and/or factor XII activation are used to treat cardiovascular, inflammatory or autoimmune disorders

Claim 5; Page 36; 43pp; English.

The present sequence represents a haemostasin polypeptide of the leech *Haementeria ghilianii*. The 14 kDa mature portion of the polypeptide is claimed. It acts as an inhibitor of the initiation complexes of blood coagulation and/or of complement. 3 variants (see also AA06449 and AA06451) of haemostasin were identified that vary at positions 16 and 60 of the mature protein. The invention provides vectors and host cells suitable for production of haemostasin. A claimed method for the treatment or prophylaxis of a condition or disorder related to C1 esterase initiation and/or Factor XII activation comprises the administration to a patient of an effective, inhibitory amount of a haemostasin polypeptide. The polypeptide may be administered by means of being expressed in the cells of the patient, the cells comprising and expressing a haemolysin nucleic acid sequence. Haemostasins can potentially be used to inhibit the activation of coagulation in the treatment of thrombotic disease such as deep venous thrombosis, pulmonary embolus, and thrombosis associated with angioplasty and endarterectomy. Disease may also be alleviated by the ability of the haemostasins to inhibit both complement activation and the intrinsic pathway of blood coagulation, such as in haemodialysis, cardiopulmonary bypass, or rejection of transplanted organs or tissues, or in cases of sepsis, myocardial infarction, stroke, particularly in the injury caused to tissues by reperfusion after an ischaemic period, atherosclerosis, shock, vasculitis, rheumatoid arthritis, sickle cell anaemia or angioedema. The haemostasins may also be used in conditions associated with activation of complement, such as: various autoimmune diseases (e.g. lupus arthritis); glomerulonephritis; nephritis; nephropathy; systemic sclerosis; Behcet's syndrome; cerebral lupus; Guillan-Barre disease; multiple sclerosis; myasthenia gravis; pemphigus; bullous pemphigoid; phototoxic reactions; thermal burns; anaphylaxis; asthma; skin reactions; infections; inflammatory bowel disease; thyroiditis; infertility; Alzheimer's disease; paroxysmal nocturnal haemoglobinuria; and haemolytic anaemia (all claimed).

XX SQ Sequence 142 AA;
 Query Match 94.9%; Score 56; DB 20; Length 142;
 Best Local Similarity 90.0%; Pred. No. 0.075;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYBYEYE 10
 |||||:
 Db 133 PNEEYDYDE 142

RESULT 13
 AAEE19950
 ID AAE19950 standard; peptide; 10 AA.
 XX AC AAE19950;
 XX DT 18-JUN-2002 (first entry)
 XX DE Complement C1s exosite binding peptide #2.
 XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW KW restenosis; myasthenia gravis.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT Modified-site 5
 FT /note= "Sulphated Phe"
 XX WO200198365-A2.
 XX PD 27-DEC-2001.
 XX PF 18-JUN-2001; 2001WO-US19405.
 XX PR 21-JUN-2000; 2000US-212998P.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI West RR, Sheppard PO, Fox BA;
 XX DR WPI; 2002-241177/29.
 XX PT New complement C1s inhibitor polypeptides for treating diseases in
 PT PT which complement activation has been shown to occur, e.g. adult
 PT PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT PT sepsis -
 XX PS Disclosure; Page 15; 99pp; English.
 XX CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC CC Molecules that inhibit complement may be used for treating diseases
 CC CC in which complement activation has been shown to occur, e.g., adult
 CC CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;
 Query Match 88.1%; Score 52; DB 23; Length 10;
 Best Local Similarity 80.0%; Pred. No. 0.017;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYBYEYE 10

```

Db      1 PNEEYDYDE 10

RESULT 14
AAE19951
ID AAE19951 standard; peptide; 10 AA.
XX
XX
AC AAE19951;
XX
XX
DT 18-JUN-2002 (first entry)
XX
XX
DE Complement C1s exosite binding peptide #3.
XX
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejections; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
XX
OS Unidentified.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 7 /note= "Sulphated Phe"
FT
XX
XX
PN WO200198365-A2.
XX
XX
PD 27-DEC-2001.
XX
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
XX
PR 21-JUN-2000; 2000US-212998P.
XX
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
XX
PI West RR, Sheppard PO, Fox BA;
XX
XX
DR WPI; 2002-241177/29.
XX
XX
PT New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
PT sepsis -
XX
XX
PS Disclosure; Page 15; 99pp; English.
XX
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejections, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
XX
SQ Sequence 10 AA;
Query Match 88.1%; Score 52; DB 23; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.017;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYDYDE 10
Db 1 PNEEYDYDE 10

Search completed: August 5, 2003, 07:38:35
Job time : 42 secs

Db      1 PNEEYDYDE 10

RESULT 15
AAE19952
ID AAE19952 standard; peptide; 10 AA.
XX
XX
AC AAE19952;
XX
XX
DT 18-JUN-2002 (first entry)
XX
XX
DE Complement C1s exosite binding peptide #4.
XX
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejections; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
XX
OS Unidentified.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 9 /note= "Sulphated Phe"
FT
XX
XX
PN WO200198365-A2.
XX
XX
PD 27-DEC-2001.
XX
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
XX
PR 21-JUN-2000; 2000US-212998P.
XX
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
XX
PI West RR, Sheppard PO, Fox BA;
XX
XX
DR WPI; 2002-241177/29.
XX
XX
PT New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
PT sepsis -
XX
XX
PS Disclosure; Page 15; 99pp; English.
XX
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejections, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
XX
SQ Sequence 10 AA;
Query Match 88.1%; Score 52; DB 23; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.017;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYDYDE 10
Db 1 PNEEYDYDE 10

Search completed: August 5, 2003, 07:38:35
Job time : 42 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:39:39 ; Search time 50 Seconds
(without alignments)
23.752 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEEYEEYE 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	10	10	US-09-883-727A-125
2	56	94.9	10	10	US-09-883-727A-2
3	56	94.9	122	10	US-09-883-727A-1
4	48	81.4	10	10	US-09-883-727A-3
5	48	81.4	10	10	US-09-883-727A-4
6	48	81.4	10	10	US-09-883-727A-5
7	48	81.4	10	10	US-09-883-727A-10
8	48	81.4	10	10	US-09-883-727A-11
9	48	81.4	10	10	US-09-883-727A-12
10	44	74.6	378	15	US-10-156-761-9424
11	41	69.5	358	16	US-10-080-170-440
12	40	67.8	10	10	US-09-883-727A-6
13	40	67.8	10	10	US-09-883-727A-7
14	40	67.8	10	10	US-09-883-727A-8
15	40	67.8	10	10	US-09-883-727A-13

16	40	67.8	10	10	US-09-883-727A-14
17	40	67.8	10	10	US-09-883-727A-15
18	40	67.8	10	10	US-09-883-727A-17
19	40	67.8	10	10	US-09-883-727A-18
20	40	67.8	10	10	US-09-883-727A-19
21	40	67.8	10	10	US-09-883-727A-20
22	40	67.8	10	10	US-09-883-727A-21
23	40	67.8	10	10	US-09-883-727A-22
24	38	64.4	193	9	US-09-815-242-13126
25	37	62.7	19	14	US-10-029-946-1
26	37	62.7	1809	9	US-09-822-635-2
27	36	61.0	283	9	US-09-764-762-5
28	36	61.0	283	10	US-09-988-975A-8
29	36	61.0	573	9	US-09-815-242-11257
30	36	61.0	773	14	US-10-067-385-8
31	36	61.0	940	10	US-09-811-088-11
32	36	61.0	940	15	US-10-314-410-11
33	35.5	60.2	10	10	US-09-883-727A-119
34	35	59.3	248	16	US-10-314-568-4
35	35	59.3	328	10	US-09-754-014-1
36	35	59.3	328	10	US-09-836-866-1
37	35	59.3	328	10	US-09-924-703-4
38	35	59.3	328	15	US-10-172-399-8
39	35	59.3	328	15	US-10-247-463-6
40	35	59.3	398	11	US-09-759-130B-445
41	35	59.3	398	14	US-10-042-431-75
42	35	59.3	439	9	US-09-883-720-6
43	35	59.3	448	16	US-10-314-568-6
44	35	59.3	503	11	US-09-920-262A-9
45	35	59.3	526	9	US-09-828-825-4

ALIGNMENTS

RESULT 1

US-09-883-727A-125
; Sequence 125, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Complement C1s inhibitor
US-09-883-727A-125

Query Match 100.0%; Score 59; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10

Db 1 PNEEYEEYE 10

RESULT 2

US-09-883-727A-2
; Sequence 2, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.

; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
US-09-883-727A-2

Query Match 94.9%; Score 56; DB 10; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0046;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
||| ||| :||
Db 1 PNEEYDYDE 10

RESULT 3
US-09-883-727A-1
; Sequence 1, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Haementaria ghilianii
US-09-883-727A-1

Query Match 94.9%; Score 56; DB 10; Length 122;
Best Local Similarity 90.0%; Pred. No. 0.064;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
||| ||| :||
Db 113 PNEEYDYDE 122

RESULT 4
US-09-883-727A-3
; Sequence 3, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-3

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
||| ||| :||
Db 1 PNEEYDYDE 10

RESULT 5
US-09-883-727A-4
; Sequence 4, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-4

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
||| ||| :||
Db 1 PNEEYDYDE 10

RESULT 6
US-09-883-727A-5
; Sequence 5, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-5

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
| | | | | : |
Db 1 PNEYEYDXE 10

RESULT 7

US-09-883-727A-10
; Sequence 10, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883.727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-10

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
| | | | | : |
Db 1 PNEYEYDXE 10

RESULT 8

US-09-883-727A-11
; Sequence 11, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883.727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)

; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-11

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
| | | | | : |
Db 1 PNEYEYDXE 10

RESULT 9

US-09-883-727A-12
; Sequence 12, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883.727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-12

Query Match 81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
| | | | | : |
Db 1 PNEYEYDXE 10

RESULT 10

US-10-156-761-9424
; Sequence 9424, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9424
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9424

Query Match 74.6%; Score 44; DB 15; Length 378;
 Best Local Similarity 80.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||
 Db 360 PEEYEEYEE 369

RESULT 11

US-10-080-170-440
 ; Sequence 440, Application US/10080170
 ; Publication No. US20030129601A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COLE, S.T.
 ; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
 ; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
 ; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
 ; FILE REFERENCE: 03495.0218
 ; CURRENT APPLICATION NUMBER: US/10/080.170
 ; CURRENT FILING DATE: 2002-06-10
 ; PRIOR APPLICATION NUMBER: 60/270,123
 ; PRIOR FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 652
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 440
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-10-080-170-440

Query Match 69.5%; Score 41; DB 16; Length 358;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EYEEYE 10
 |||||
 Db 181 EYEEYE 187

RESULT 12

US-09-883-727A-6
 ; Sequence 6, Application US/09883727A
 ; Patent No. US20020102256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: West, Robert R.
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Fox, Brian
 ; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
 ; TITLE OF INVENTION: Complement C1s
 ; FILE REFERENCE: 00-33
 ; CURRENT APPLICATION NUMBER: US/09/883,727A
 ; CURRENT FILING DATE: 2001-09-18
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: C1s exosite binding moiety
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (5)...(5)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (7)...(7)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 US-09-883-727A-6

Query Match 67.8%; Score 40; DB 10; Length 10;
 Best Local Similarity 70.0%; Pred. No. 1.5;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||
 Db 1 PNEEXEXDYE 10

RESULT 13

US-09-883-727A-7
 ; Sequence 7, Application US/09883727A
 ; Patent No. US20020102256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: West, Robert R.
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Fox, Brian
 ; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
 ; TITLE OF INVENTION: Complement C1s
 ; FILE REFERENCE: 00-33
 ; CURRENT APPLICATION NUMBER: US/09/883,727A
 ; CURRENT FILING DATE: 2001-09-18
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: C1s exosite binding moiety
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (5)...(5)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (9)...(9)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 US-09-883-727A-7

Query Match 67.8%; Score 40; DB 10; Length 10;
 Best Local Similarity 70.0%; Pred. No. 1.5;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||
 Db 1 PNEEXEXDYE 10

RESULT 14

US-09-883-727A-8
 ; Sequence 8, Application US/09883727A
 ; Patent No. US20020102256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: West, Robert R.
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Fox, Brian
 ; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
 ; TITLE OF INVENTION: Complement C1s
 ; FILE REFERENCE: 00-33
 ; CURRENT APPLICATION NUMBER: US/09/883,727A
 ; CURRENT FILING DATE: 2001-09-18
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: C1s exosite binding moiety
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (7)...(7)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 ; NAME/KEY: MUTAGEN
 ; LOCATION: (9)...(9)
 ; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
 US-09-883-727A-8

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
||||| :
Db 1 PNEYEXDXE 10

RESULT 15

US-09-883-727A-13
; Sequence 13, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Fox, Brian
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883.727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = sulfated tyrosine
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-13

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
||||| :
Db 1 PNEYEXDXE 10

Search completed: August 5, 2003, 07:48:01
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:37:55 ; Search time 17 Seconds
(without alignments)
24.889 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYEE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	66.1	307	4 US-09-107-532A-6172	Sequence 6172, Ap
2	39	66.1	534	4 US-09-107-532A-6549	Sequence 6549, Ap
3	37	62.7	19	4 US-10-029-946-1	Sequence 1, Appli
4	37	62.7	635	4 US-09-328-352-8029	Sequence 8029, Ap
5	37	62.7	2183	3 US-08-746-111-5	Sequence 5, Appli
6	36	61.0	172	4 US-08-793-273C-29	Sequence 29, Appli
7	36	61.0	263	2 US-08-790-137-4	Sequence 4, Appli
8	36	61.0	263	2 US-08-824-874-5	Sequence 5, Appli
9	36	61.0	263	3 US-08-807-151-5	Sequence 5, Appli
10	36	61.0	263	3 US-09-210-084-5	Sequence 5, Appli
11	36	61.0	263	4 US-09-478-957-5	Sequence 5, Appli
12	36	61.0	263	4 US-09-764-762-5	Sequence 5, Appli
13	36	61.0	368	4 US-09-724-519-10	Sequence 10, Appli
14	36	61.0	368	4 US-09-592-037-10	Sequence 10, Appli
15	36	61.0	382	4 US-09-724-519-4	Sequence 4, Appli
16	36	61.0	382	4 US-09-592-037-4	Sequence 4, Appli
17	36	61.0	404	4 US-09-134-001C-4848	Sequence 4848, Ap
18	36	61.0	513	4 US-09-724-519-6	Sequence 6, Appli
19	36	61.0	513	4 US-09-592-037-6	Sequence 6, Appli
20	36	61.0	575	4 US-09-724-519-8	Sequence 8, Appli
21	36	61.0	575	4 US-09-592-037-8	Sequence 8, Appli
22	36	61.0	940	2 US-08-938-365-4	Sequence 4, Appli
23	36	61.0	941	1 US-08-343-760A-2	Sequence 2, Appli
24	36	61.0	1053	4 US-09-724-519-2	Sequence 2, Appli
25	36	61.0	1053	4 US-09-592-037-2	Sequence 2, Appli
26	36	61.0	1056	4 US-09-595-684B-29	Sequence 29, Appli
27	36	61.0	1057	3 US-09-541-782-10	Sequence 10, Appli

28	36	61.0	1057	4 US-09-723-820-10	Sequence 10, Appli
29	36	61.0	1810	4 PCT-US95-273C-4	Sequence 4, Appli
30	36	61.0	1810	5 PCT-US95-11684-4	Sequence 4, Appli
31	35.5	60.2	17	3 US-09-150-133-23	Sequence 23, Appli
32	35.5	60.2	17	3 US-09-150-141-23	Sequence 23, Appli
33	35.5	60.2	17	3 US-09-374-493-23	Sequence 23, Appli
34	35.5	60.2	17	3 US-09-374-824-23	Sequence 23, Appli
35	35.5	60.2	17	3 US-09-374-492-23	Sequence 23, Appli
36	35	59.3	264	4 US-09-134-001C-5035	Sequence 5035, Ap
37	35	59.3	267	4 US-09-134-001C-4539	Sequence 4539, Ap
38	35	59.3	306	1 US-08-424-682A-1	Sequence 1, Appli
39	35	59.3	320	4 US-09-134-001C-4204	Sequence 4204, Ap
40	35	59.3	328	1 US-08-285-087-2	Sequence 2, Appli
41	35	59.3	328	1 US-08-186-529-2	Sequence 2, Appli
42	35	59.3	328	1 US-08-621-493-2	Sequence 2, Appli
43	35	59.3	328	1 US-08-640-386A-2	Sequence 2, Appli
44	35	59.3	328	2 US-08-965-688-2	Sequence 2, Appli
45	35	59.3	328	2 US-08-751-767A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-107-532A-6172

; Sequence 6172, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6172:

SEQUENCE CHARACTERISTICS:

LENGTH: 307 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...307

SEQUENCE DESCRIPTION: SEQ ID NO: 6172:

US-09-107-532A-6172

Query Match 66.1%; Score 39; DB 4; Length 307;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEYEYEEYE 10
DB 251 PNEYEROYK 260
|||||:|:

RESULT 2
US-09-107-532A-6549
; Sequence 6549, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arintello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6549:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...534
; SEQUENCE DESCRIPTION: SEQ ID NO: 6549:
US-09-107-532A-6549

Query Match 66.1%; Score 39; DB 4; Length 534;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PNEYEYEEYE 10
DB 472 PTNEYKYSYE 481
|||||:

RESULT 3
US-10-029-946-1
; Sequence 1, Application US/10029946

Patent No. 6579897
; GENERAL INFORMATION:
; APPLICANT: TANG, PENG CHO
; APPLICANT: SUN, LI
; APPLICANT: MCMAHON, GERALD
; TITLE OF INVENTION: 3-(CYCLOALKANOHEXTEROARYLIDENYL)-2-INDOLINONE PROTEIN
; TITLE OF INVENTION: TYROSINE KINASE INHIBITORS
; FILE REFERENCE: 038602/1258
; CURRENT APPLICATION NUMBER: US/10/029,946
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/482,198
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 09/099,721
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/050,413
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: 60/059,544
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059,384
; PRIOR FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; US-10-029-946-1

Query Match 62.7%; Score 37; DB 4; Length 19;
Best Local Similarity 87.5%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EYEEYEEYE 10
DB 5 EYEEYEEYE 12
|||||

RESULT 4
US-09-328-352-8029
; Sequence 8029, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8029
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-8029

Query Match 62.7%; Score 37; DB 4; Length 635;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEYEYEEYE 10
DB 106 PNKTYEYKFK 115
|||:||||:

RESULT 5
US-08-746-111-5
; Sequence 5, Application US/08746111
; Patent No. 6066778
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David

APPLICANT: Cui, Jisong
TITLE OF INVENTION: Compositions And Methods For Screening
; Title Of Invention: Compounds For Anticoagulant Activity
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,111
FILING DATE: 06-NOV-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02536
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2183 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-746-111-5

Query Match 62.7%; Score 37; DB 3; Length 2183;
Best Local Similarity 66.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEY 9
Db 143 PGEYTYEW 151

RESULT 6
US-08-793-273C-29
; Sequence 29, Application US/08793273C
; Patent No. 6482410
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Phillips, Greg
; APPLICANT: Prieto, Anne L.
; TITLE OF INVENTION: CYTOTOXIC DERIVATIVES THAT STIMULATE ATTACHMENT AND
; Title Of Invention: NEURITE OUTGROWTH, AND METHODS OF MAKING SAME
FILE REFERENCE: BSC00225
CURRENT APPLICATION NUMBER: US/08/793,273C
CURRENT FILING DATE: 1997-05-22
PRIOR APPLICATION NUMBER: PCT/US95/11684
PRIOR FILING DATE: 1995-09-14
PRIOR APPLICATION NUMBER: 08/308,359
PRIOR FILING DATE: 1994-09-16
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 29
LENGTH: 172
TYPE: PRT
ORGANISM: Gallus gallus
US-08-793-273C-29

Query Match 61.0%; Score 36; DB 4; Length 172;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEY 8
Db 63 PTEYEEY 70
RESULT 7
US-08-790-137-4
; Sequence 4, Application US/08790137
; Patent No. 5840871
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED
; Title Of Invention: KALLIKREIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/790,137
FILING DATE: Filed Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0195 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 55527
US-08-790-137-4

Query Match 61.0%; Score 36; DB 2; Length 263;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEY 10
Db 163 PTEFEYSHD 172

RESULT 8
US-08-824-874-5
; Sequence 5, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; Number Of Sequences: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive

```

; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 55527
; US-08-824-874-5

Query Match 61.0%; Score 36; DB 2; Length 263;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEVEYEYE 10
Db 163 PTEEFYSHD 172

RESULT 9
US-08-807-151-5
; Sequence 5, Application US/08807151
; Patent No. 6043033
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL HUMAN PROSTATE-ASSOCIATED
; TITLE OF INVENTION: PROTEASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,151
; FILING DATE: Filed Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0227 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 55527
; US-08-807-151-5

Query Match 61.0%; Score 36; DB 3; Length 263;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEVEYEYE 10
Db 163 PTEEFYSHD 172

RESULT 10
US-09-210-084-5
; Sequence 5, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 55527
; US-09-210-084-5

Query Match 61.0%; Score 36; DB 3; Length 263;

```

Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
Db 163 PTEEFYSHD 172

RESULT 11

US-09-478-957-5
; Sequence 5, Application US/09478957
; Patent No. 6350448

GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL HUMAN PROSTATE-ASSOCIATED
; TITLE OF INVENTION: PROTEASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US

; ZIP: 94304

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA: US/09/478,957
; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA: 08/807,151
; FILING DATE:

ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0227 US
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

; LENGTH: 263 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 55527

US-09-478-957-5

Query Match 61.0%; Score 36; DB 4; Length 263;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
Db 163 PTEEFYSHD 172

RESULT 12

US-09-764-762-5
; Sequence 5, Application US/09764762
; Patent No. 6472195

GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: NOVEL KALLIKREIN

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

; ZIP: 94304

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA: US/09/764,762

; FILING DATE: 16-Jan-2001

; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/210,084

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0252 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

; LENGTH: 263 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 55527

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-764-762-5

Query Match 61.0%; Score 36; DB 4; Length 263;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYEEYE 10
Db 163 PTEEFYSHD 172

RESULT 13

US-09-724-519-10

; Sequence 10, Application US/09724519

; Patent No. 6414121

GENERAL INFORMATION:

; APPLICANT: Wood, Kenneth

; APPLICANT: Finer, Jeffrey

; APPLICANT: Beraud, Christophe

; APPLICANT: Mak, John

; APPLICANT: Sakowicz, Roman

; TITLE OF INVENTION: Methods of screening for modulators of

; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell

; TITLE OF INVENTION: proliferation states

; FILE REFERENCE: 1014A

; CURRENT APPLICATION NUMBER: US/09/724,519

; CURRENT FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 09/592,037

; PRIOR FILING DATE: 2000-06-12

; PRIOR APPLICATION NUMBER: 09/428,156

; PRIOR FILING DATE: 1999-10-27

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 368

; TYPE: PRT

; ORGANISM: Human

US-09-724-519-10

Query Match 61.0%; Score 36; DB 4; Length 368;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYE 8
||||| :
Db 121 PNEYTWE 128

RESULT 14

US-09-592-037-10
; Sequence 10, Application US/09592037
; Patent No. 6437115
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/592,037
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Human
US-09-592-037-10

Query Match 61.0%; Score 36; DB 4; Length 368;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYE 8
||||| :
Db 121 PNEYTWE 128

RESULT 15

US-09-724-519-4
; Sequence 4, Application US/09724519
; Patent No. 6414121
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-4

Query Match 61.0%; Score 36; DB 4; Length 382;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEYE 8
||||| :
Db 121 PNEYTWE 128

Search completed: August 5, 2003, 07:40:18
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:47:10 ; Search time 39 Seconds
(without alignments)
24.659 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEEYEVYE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	40.7	9	2 S70345	amine oxidase (cop
2	24	40.7	10	2 S30348	clotting protein -
3	21	35.6	6	2 S11556	hydrogenulfite re
4	20	33.9	9	2 D41978	callifMRPamide 4 -
5	19	32.2	7	2 A38081	amine oxidase (cop
6	19	32.2	9	2 S70332	endosperm protein,
7	18	30.5	9	2 A44787	callifMRPamide 10
8	17	28.8	8	2 S29272	tocopherol-binding
9	17	28.8	9	2 PH0108	late G1-69 protein
10	17	28.8	10	2 A43405	6-phosphofructo-2-
11	16	27.1	7	4 I56695	hypothetical I2 pr
12	16	27.1	9	2 C24180	fibrinogen beta ch
13	16	27.1	9	2 D24180	fibrinogen beta ch
14	16	27.1	9	2 PT0299	Ig heavy chain CRD
15	16	27.1	9	2 S36850	Ig heavy chain V r
16	16	27.1	10	1 RHLWGS	gonadoliberin - se
17	16	27.1	10	2 S62208	polyferredoxin - M
18	15	25.4	5	2 S68326	blood cell protein
19	15	25.4	9	2 S30494	cat gene leader pe
20	15	25.4	9	2 B24362	chloramphenicol O-
21	15	25.4	9	2 B41978	callifMRPamide 2 -
22	15	25.4	9	2 C41978	callifMRPamide 3 -
23	15	25.4	10	2 S66635	alpha-2-macroglobu
24	15	25.4	10	2 S38304	lectin GNL1 alpha
25	15	25.4	10	2 S62880	polygalacturonase
26	14	23.7	7	1 A61324	dermorphin - Rohde
27	14	23.7	7	2 A60139	fatty-acid synthas
28	14	23.7	7	2 S36662	dermorphin (lys-7)
29	14	23.7	7	2 A44428	platelet aggregati

30 14 23.7 8 2 PT0030 inulinase (EC 3.2.
31 14 23.7 8 2 S68325 blood cell protein
32 14 23.7 8 2 I64832 Ca2+-transporting
33 14 23.7 9 2 A61357 phyllocaerulein -
34 14 23.7 9 2 A41978 callifMRPamide 1 -
35 14 23.7 9 2 E41978 callifMRPamide 5 -
36 14 23.7 10 2 S39392 calpain (EC 3.4.22
37 14 23.7 10 2 A49187 gonadotropin-relea
38 14 23.7 10 2 PT0291 Ig heavy chain CRD
39 14 23.7 10 2 PT0215 T-cell receptor be
40 13 22.0 4 2 S55238 pallidipin - aspas
41 13 22.0 5 2 A32014 tram protein - Esc
42 13 22.0 6 2 A44916 mosquitoicidal toxi
43 13 22.0 6 2 S29637 jacalin beta-II ch
44 13 22.0 7 2 S21230 dermorphin (Trp-4,
45 13 22.0 7 2 A15398 choline oxidase (E

ALIGNMENTS

RESULT 1

S70345

amine oxidase (copper-containing) (EC 1.4.3.6) II - Aspergillus niger (fragments)
C;Species: Aspergillus niger
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S70345

R;Rebort, I.; Pec, P.; Luhova, L.; Toyama, H.; Matsushita, K.; Hirota, S.; Kitagawa, T

Biochim. Biophys. Acta 1295, 59-72, 1996

A;Title: Two amine oxidases from Aspergillus niger AKU 3302 contain topa quinone as the

A;Reference number: S70344; MUID:96283794; PMID:8679675

A;Accession: S70345

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-5-6-9 <PRE>

C;Keywords: oxidoreductase

Query Match 40.7%; Score 24; DB 2; Length 9;

Best Local Similarity 80.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEY 5

Db 5 PNKEY 9

RESULT 2

S30348

clotting protein - signal crayfish
C;Species: Pacifastacus leniusculus (signal crayfish)

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999

C;Accession: S30348

R;Kopacek, P.; Hall, M.; Soederhaell, K.

Eur. J. Biochem. 213, 591-597, 1993

A;Title: Characterization of a clotting protein, isolated from plasma of the freshwater

A;Reference number: S30348; MUID:93238739; PMID:8097463

A;Accession: S30348

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <KOP>

Query Match 40.7%; Score 24; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NEEYEV 7

Db 4 NLEYQY 9

RESULT 3

S11556

hydrogensulfite reductase (EC 1.8.99.3) chain 2 - Desulfovibrio thermophilus (fragment)

N:Alternate names: bisulfite reductase; desulfofusicidin
C:Species: Desulfovibrio thermophilus
C:Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C:Accession: S11556
R:Fauque, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.; L
Biochim. Biophys. Acta 1040, 112-118, 1990
A:Title: Purification and characterization of bisulfite reductase (desulfofusicidin) from
A:Reference number: S11024; MUID:90335276; PMID:2165817
A:Accession: S11556
A:Molecule type: protein
A:Residues: 1-6 <FAU>
C:Keywords: oxidoreductase

Query Match 35.6%; Score 21; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEYEY 6
| | | |
DB 1 PEEKYK 6

RESULT 4

D41978
calliFMRamide 4 - bluebottle fly (Calliphora vomitoria)
C:Species: Calliphora vomitoria
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C:Accession: D41978
R:Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A:Reference number: A41978; MUID:92196111; PMID:1549595
A:Accession: D41978
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <DUV>
C:Keywords: amidated carboxyl end; neuropeptide
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 33.9%; Score 20; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEYEY 5
| | | |
DB 2 PNQDF 6

RESULT 5

A38081
amine oxidase (copper-containing) (EC 1.4.3.6) - yeast (Pichia angusta) (fragment)
C:Species: Pichia angusta
C:Date: 31-Dec-1993 #sequence_revision 03-Feb-1994 #text_change 20-Apr-2000
C:Accession: A38081
R:Mu, D.; Jones, S.M.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Klinman, J.P.
J. Biol. Chem. 267, 7979-7982, 1992
A:Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine o
A:Reference number: A38081; MUID:92235001; PMID:1589055
A:Accession: A38081
A:Molecule type: protein
A:Residues: 1-7 <NUA>
C:Keywords: copper; oxidoreductase; quinoprotein; topaquinone
F:4/Modified site: topaquinone (Tyr) #status experimental

Query Match 32.2%; Score 19; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YEY 7
| | | |
DB 4 YEY 6

RESULT 6

S70332
endosperm protein, 10K - rye (fragment)
C:Species: Secale cereale (rye)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C:Accession: S70332
R:Roche, A.; Calero, M.; Soriano, F.; Mendez, E.
Biochim. Biophys. Acta 1295, 13-22, 1996
A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.
A:Reference number: S70327; MUID:96283789; PMID:8679669
A:Accession: S70332
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <ROC>

Query Match 32.2%; Score 19; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YEY 7
| | | |
DB 4 YEY 6

RESULT 7

A44787
calliFMRamide 10 - bluebottle fly (Calliphora vomitoria)
C:Species: Calliphora vomitoria
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C:Accession: A44787
R:Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorp
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (des
A:Reference number: A41978; MUID:92196111; PMID:1549595
A:Accession: A44787
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <DUV>
C:Keywords: amidated carboxyl end; neuropeptide
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 30.5%; Score 18; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEYEY 5
| | | |
DB 2 PNQDF 6

RESULT 8

S29272
tocopherol-binding protein, 81K - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
C:Accession: S29272
R:Nalecz, K.A.; Nalecz, M.J.; Azzi, A.
Eur. J. Biochem. 209, 37-42, 1992
A:Title: Isolation of tocopherol-binding proteins from the cytosol of smooth muscle A7r
A:Reference number: S29272; MUID:93011150; PMID:1396710
A:Accession: S29272
A:Molecule type: protein
A:Residues: 1-8 <NAL>
A:Experimental source: smooth muscle A7r5 cells

Query Match 28.8%; Score 17; DB 2; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 BEVEY 9
| | | |
DB 2 EEDQY 8

RESULT 9

PH0108
late G1-69 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH0108
R;Nikaido, T.; Bradley, D.W.; Pardee, A.B.
Exp. Cell Res. 192, 102-109, 1991
A;Title: Molecular cloning of transcripts that accumulate during the late G1 phase in culture
A;Reference number: PH0108; MUID:91078351; PMID:1984406
A;Accession: PH0108
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-9 <NIK>

Query Match 28.8%; Score 17; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EY 5
|||
Db 7 EY 9

RESULT 10

A43405
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC 3.1.3.4)
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Mar-2000
C;Accession: A43405
R;Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilakis, S.J.; Bartrons, R.
J. Biol. Chem. 267, 17939-17943, 1992
A;Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence for a bifunctional enzyme
A;Reference number: A43405; MUID:92388154; PMID:1325453
A;Accession: A43405
A;Molecule type: protein
A;Residues: 1-10 <VEN>
C;Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phosphoglycerate kinase; phosphoric monoester hydrolase; phosphotransferase

Query Match 28.8%; Score 17; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 3.3e+03;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NEEYEV 9
| : : ||
Db 3 NQHLKGY 10

RESULT 11

I56695
hypothetical L2 protein (mistranslated) - human papillomavirus type 16 (fragment)
C;Species: human papillomavirus type 16
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C;Accession: I56695
R;Schneider-Maunoury, S.; Croissant, O.; Orth, G.
J. Virol. 61, 3295-3298, 1987
A;Title: Integration of human papillomavirus type 16 DNA sequences: a possible early event
A;Reference number: I56695; MUID:87311896; PMID:3041049
A;Accession: I56695
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-7 <SCH>
A;Cross-references: GB:M30709; NID:9190253; PIDN:AAA65995.1; PID:9553616
C;Comment: This is the hypothetical translation of a viral sequence integrated into the genome. It is translated in an incorrect, -1, reading frame of the L2 protein.

Query Match 27.1%; Score 16; DB 4; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 YEYEV 9

Db 1 HSYRY 5

RESULT 12

C24180
fibrinogen beta chain - Japanese macaque (fragment)
N;Contains: fibrinopeptide B
C;Species: Macaca fuscata (Japanese macaque)
C;Date: 05-Jun-1988 #sequence_revision 10-Mar-1994 #text_change 26-Jan-1996
C;Accession: C24180
R;Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 97, 1487-1492, 1985
A;Title: Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and patas monkey (Ptilinops) and baboons.
A;Reference number: A91990; MUID:85289140; PMID:3928610
A;Accession: C24180
A;Molecule type: protein
A;Residues: 1-9 <NAK>
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide isomerase

Query Match 27.1%; Score 16; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEE 4
|||
Db 1 NEE 3

RESULT 13

D24180
fibrinogen beta chain - red guenon (fragment)
N;Contains: fibrinopeptide B
C;Species: Erythrocebus patas (red guenon, hussar)
C;Date: 05-Jun-1988 #sequence_revision 10-Mar-1994 #text_change 26-Jan-1996
C;Accession: D24180
R;Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 97, 1487-1492, 1985
A;Title: Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and patas monkey (Ptilinops) and baboons.
A;Reference number: A91990; MUID:85289140; PMID:3928610
A;Accession: D24180
A;Molecule type: protein
A;Residues: 1-9 <NAK>
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide isomerase

Query Match 27.1%; Score 16; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEE 4
|||
Db 1 NEE 3

RESULT 14

PT0299
Ig heavy chain CDR3 region (clone 5-103B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0299
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and hypervariability in the CDR3 region of the Ig heavy chain CDR3 region (clone 5-103B) - human (fragment)
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0299
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotrimer; immunoglobulin

Query Match 27.1%; Score 16; DB 2; Length 9;

Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 EYEV 7
| : |
Db 2 EWDY 6

RESULT 15

S36850
Ig heavy chain V region - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1994 #sequence_revision 01-Dec-1995 #text_change 05-Nov-1999
C;Accession: S36850
R;Jacob, J.; Kelsoe, G.
submitted to the EMBL Data Library, July 1992
A;Description: In situ studies on the primary immune response to (4-hydroxy-3-nitrophenyl)
A;Reference number: S25024
A;Accession: S36850
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-9 <JAC>
A;Cross-references: EMBL:X67387; NID:G50113; PID:CAA47799.1; PID:e51594; PID:gl333871
C;Keywords: heterotetramer; immunoglobulin

Query Match 27.1%; Score 16; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YEV 7
| : |
Db 3 YDY 5

Search completed: August 5, 2003, 07:52:26
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:40:25 ; Search time 23 Seconds
(without alignments)
20.446 Million cell updates/sec

Title: US-09-883-727A-125
Perfect score: 59
Sequence: 1 PNEYEYEEYE 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	33.9	9	1 FAR4_CALVO	P41859 calliphora
2	20	33.9	10	1 UH05_RAT	P56573 rattus norv
3	18	30.5	9	1 FAR4_CALVO	P41865 calliphora
4	17	28.8	10	1 AL19_CARMA	P81822 carcinus ma
5	16	27.1	9	1 FIBB_ERYPA	P19346 erythrocebu
6	16	27.1	9	1 FIBB_MACFU	P19345 macaca fusc
7	16	27.1	10	1 GON1_PETMA	P04378 petromyzon
8	15	25.4	7	1 FAR2_ASCSU	P31890 ascaris suu
9	15	25.4	7	1 UF03_MOUSE	P38641 mus musculu
10	15	25.4	8	1 PLP_BRANA	P81707 brassica na
11	15	25.4	8	1 UC26_MAIZE	P80632 zea mays (m
12	15	25.4	9	1 FAR2_CALVO	P41857 calliphora
13	15	25.4	9	1 FAR3_CALVO	P41858 calliphora
14	15	25.4	9	1 FLA2_TREHY	P80159 treponema h
15	15	25.4	9	1 FRF1_SARBU	P83350 sarcophaga
16	15	25.4	9	1 LFCA_STAAU	P36884 staphylococ
17	14	23.7	7	1 E105_LITRU	P82101 litorea rub
18	14	23.7	8	1 AL16_CARMA	P81819 carcinus ma
19	14	23.7	9	1 FAR1_CALVO	P41856 calliphora
20	14	23.7	9	1 FAR5_CALVO	P41860 calliphora
21	14	23.7	9	1 UPAT7_HUMAN	P30093 homo sapien
22	13	22.0	5	1 TRM3_ECOLI	P13973 escherichia
23	13	22.0	6	1 ASP2_LACSN	P82655 lactobacill
24	13	22.0	7	1 ALL2_CARMA	P81805 carcinus ma
25	13	22.0	7	1 ALL3_CARMA	P81806 carcinus ma
26	13	22.0	7	1 CHOX_ALCSP	P16101 alcaligenes
27	13	22.0	7	1 FAR4_PANRE	P41875 panagrellus
28	13	22.0	8	1 FAR1_PANRE	P41872 panagrellus
29	13	22.0	8	1 HTF1_PERAM	P04548 periplaneta
30	13	22.0	8	1 HTF2_PERAM	P04549 periplaneta
31	13	22.0	8	1 HTF_TENNO	P25419 tenebrio mo
32	13	22.0	9	1 FAR2_PANRE	P41873 panagrellus
33	13	22.0	9	1 FIBB_PAPAN	P19344 papio anubi

ALIGNMENTS

RESULT 1

FAR4_CALVO
ID FAR4_CALVO STANDARD; PRT; 9 AA.
AC P41859;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliMRPamide 4.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Dave H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Renfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliMRPamides) from the blowfly
RT Calliphora vomitoria";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; D41978; D41978.
KW Neuropeptide; Amidation.
FT MOD RES 9 9
SQ SEQUENCE 9 AA; 1182 MW; 31730699CAB6D457 CRC64;
AMIDATION.

Query Match 33.9%; Score 20; DB 1; Length 9;
Best Local Similarity 40.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEYEY 5
Db |||::: 6
2 PNQDF 6

RESULT 2

UH05_RAT
ID UH05_RAT STANDARD; PRT; 10 AA.
AC P56573;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of heart tissue (Spot P5) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleisner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;

P19343 papio hamad
P19342 theropithec
P81179 diaprepes a
P81731 helicoverpa
P30426 bothrops in
P81110 romalea mic
P11385 carausius m
P56923 rana tempor
P08610 phyllomedus
P30090 homo sapien
P82158 cydia pomon
P81820 carcinus ma

34 13 22.0 9 1 FIBB PAPH
35 13 22.0 9 1 FIBB THEGE
36 13 22.0 9 1 PGLR DIAAB
37 13 22.0 10 1 AMPN HELAM
38 13 22.0 10 1 BPP8 BOTIN
39 13 22.0 10 1 HTF1 ROMMI
40 13 22.0 10 1 HTF2 CARMO
41 13 22.0 10 1 TEMK RANTE
42 13 22.0 10 1 TKN PHYBI
43 13 22.0 10 1 UPA4 HUMAN
44 12 20.3 7 1 ALL7 CYDPO
45 12 20.3 8 1 AL17 CARMA

RL Submitted (SEP-1998) to the SWISS-PROT data bank.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 8.3, ITS MW IS: 30 kDa.
 FT UNSURE 9 9 OR P.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1269 MW; 00CBA4BB46C5BAB8 CRC64;

Query Match 33.9%; Score 20; DB 1; Length 10;
 Best Local Similarity 28.6%; Pred. No. 8e+02;
 Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EYEYEE 10
 : : :
 Db 2 QYDSQYD 8

RESULT 3-

ID FARA_CALVO STANDARD; PRT; 9 AA.
 AC P41865;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CalliFMRamide 10.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Thoracic ganglion;
 RX MEDLINE=92196111; PubMed=1549595;
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A.;
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
 RT neuropeptides (designated calliFMRamides) from the blowfly
 RT Calliphora vomitoria.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; A44787;
 DR Neuropeptide; Amidation.
 KW MOD RES 9
 FT UNSURE 1 1
 FT OR S OR A.
 SQ SEQUENCE 9 AA; 1183 MW; 29D00699CAB40457 CRC64;

Query Match 30.5%; Score 18; DB 1; Length 9;
 Best Local Similarity 40.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEYEY 5
 : : :
 Db 2 PNRDF 6

RESULT 4

ID ALI9_CARMA STANDARD; PRT; 10 AA.
 AC P81822;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 19.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Pleocyemata; Brachyura;
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD RES 10 10
 FT AMIDATION (POTENTIAL).
 SQ SEQUENCE 10 AA; 1101 MW; 96687CD5AB569AB1 CRC64;

Query Match 28.8%; Score 17; DB 1; Length 10;
 Best Local Similarity 28.6%; Pred. No. 2.4e+03;
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PNEYEY 7
 : : :
 Db 2 PTDMYSF 8

RESULT 5

ID FIBB_ERYPA STANDARD; PRT; 9 AA.
 AC P19346;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Erythrocybus patas (Red guenon) (Hussar).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Erythrocybus.
 OX NCBI_TaxID=9538;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85289140; PubMed=3928610;
 RA Nakamura S., Takenaka O., Takahashi K.;
 RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
 RT patas monkey (Erythrocybus patas): their amino acid sequences,
 RT restricted mutations, and a molecular phylogeny for macaques,
 RT guenons, and baboons.";
 RL J. Biochem. 97:1487-1492(1985).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR PIR; D24180;
 DR InterPro; IPR002181; Fibrinogen C.
 DR PROSITE; PS00514; FIBRIN AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9
 FT NON_TER 9 9
 FT FIBRINOPEPTIDE B.
 SQ SEQUENCE 9 AA; 1020 MW; 69FE7879C732CB1B CRC64;

Query Match 27.1%; Score 16; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEE 4
 : : :
 Db 1 NEE 3

RESULT 6

FIBB_MACFU
 ID FIBB_MACFU STANDARD; PRT; 9 AA.
 AC P19345;

DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain (Contains: Fibrinopeptide B) (Fragment).
 GN FGB.
 OS Macaca fuscata fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OC NCBI_TaxID=9543;
 RN [1]
 RN SEQUENCE.
 RP MEDLINE=85289140; PubMed=3928610;
 RX Nakamura S., Takenaka O., Takahashi K.;
 RA "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
 RT patas monkey (Erythrocebus patas): their amino acid sequences,
 RT restricted mutations, and a molecular phylogeny for macaques,
 RT guenons, and baboons";
 RL J. Biochem. 97:1487-1492 (1985).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 CC PIR; C24180;
 DR InterPro: IPR002181; Fibrinogen C.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON TER 9 9
 SQ SEQUENCE 9 AA; 1038 MW; 69FE65BC735BB1B CRC64;

 Query Match 27.1%; Score 16; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 NEE 4
 DB 1 NEE 3

 RESULT 7
 ID_GONI1_PETMA STANDARD; PRT; 10 AA.
 AC P04376;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I)
 DE (Luliberin I).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OC NCBI_TaxID=7757;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Brain;
 RX Sherwood N.M., Sower S.A., Marshak D.R., Fraser B.A., Brownstein M.J.;
 RA "Primary structure of gonadotropin-releasing hormone from lamprey
 RT brain";
 RL J. Biol. Chem. 261:4812-4819 (1986).
 CC -1- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the GnRH family.
 CC PIR; A01412; RHLNMG.
 DR InterPro: IPR002012; GnRH.

DR Pfam; PF00446; GnRH; 1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION
 SQ SEQUENCE 10 AA; 1244 MW; 1E4B36237B1735AB CRC64;

 Query Match 27.1%; Score 16; DB 1; Length 10;
 Best Local Similarity 25.0%; Pred. No. 3.4e+03;
 Matches 2; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 QY 3 EYEYEYE 10
 DB 1 QHYSLEWK 8

 RESULT 8
 ID_FAR2_ASCSU STANDARD; PRT; 7 AA.
 AC P31890;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE FMRamide-like neuropeptide AP2.
 OS Ascaris suum (Big roundworm) (Ascaris lumbricoides), and
 OS Panagrellus redivivus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridoidea;
 OC Ascarididae; Ascaris.
 OC NCBI_TaxID=6253, 6233;
 RN [1]
 RN SEQUENCE.
 RC SPECIES=A.suum;
 RX MEDLINE=93324431; PubMed=8332542;
 RA Cowden C., Stretton A.O.W.;
 RT "AP2, an Ascaris neuropeptide: isolation, sequence, and bioactivity";
 RL Peptides 14:423-430 (1993).
 RN [2]
 RC SPECIES=P.redivivus;
 RX MEDLINE=95060998; PubMed=7970891;
 RA Maule A.G., Shaw C., Bowman J.W.;
 RT "The FMRamide-like neuropeptide AP2 (Ascaris suum) is present in the
 RT free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
 RL Parasitology 109:351-356 (1994).
 CC -1- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF
 CC GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 7 7 AMIDATION.
 SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B11E350 CRC64;

 Query Match 25.4%; Score 15; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 QY 4 EYE 7
 DB 1 KHEY 4

 RESULT 9
 ID_UF03_MOUSE STANDARD; PRT; 7 AA.
 AC P38641;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P36) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.1, ITS MW IS: 36 Kda.
FT NON TER 7
SQ SEQUENCE 7 AA; 842 MW; 6AA72B1DB1B1180 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 7;
Best Local Similarity 42.9%; Pred. No. 1.3e+05;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 NEEYEY 8
:|:|:
Db 1 HEEAELD 7

RESULT 10
PLP BRANA
ID - PLP BRANA STANDARD; PRT; 8 AA.
AC P81707;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Plastidial lipid-associated protein (Fragment).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. TOPAZ; TISSUE=Tapetum;
RX MEDLINE=9349136; PubMed=10420651;
RA Hernandez-Pinzon I., Ross J.H.E., Barnes K.A., Damant A.P.,
RA Murphy D.J.;
RT "Composition and role of tapetal lipid bodies in the biogenesis of the
RT pollen coat of Brassica napus.";
RL Planta 208:588-598(1999).
CC -1- FUNCTION: MAY PLAY A STRUCTURAL ROLE IN THE ELAIOTPLAST, A TAPETUM-
CC SPECIFIC PLASTIDIAL LIPID ORGANELLE.
CC -1- TISSUE SPECIFICITY: TAPETUM OF ANTHEERS.
FT NON TER 8
SQ SEQUENCE 8 AA; 989 MW; 9D7B1AA452CAA042 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NEEY 5
:|:
Db 5 NDEW 8

RESULT 11
UC26 MAIZE
ID - UC26 MAIZE STANDARD; PRT; 8 AA.
AC P80632;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 907)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Trouzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 7.0, ITS MW IS: 57.2 Kda.
DR Maize-2DPAGE; P80632; COLEOPTILE.
FT NON TER 1
FT NON TER 8
SQ SEQUENCE 8 AA; 990 MW; 9639D6DAB4176B1D CRC64;

Query Match 25.4%; Score 15; DB 1; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.3e+05;
Matches 1; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYE 6
:|:
Db 3 PRDQFX 8

RESULT 12
FAR2 CALVO
ID - FAR2 CALVO STANDARD; PRT; 9 AA.
AC P41857;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRamide 2.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnson A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
CC SALIVARY GLAND OF CALLIPHORA.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; B41978; B41978.
KW Neuropeptide; Amidation.
FT MOD RES 9
FT SEQUENCE 9 AA; 1128 MW; 29D00699CAB6C5A7 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 9;
Best Local Similarity 20.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEY 5
:|:
Db 2 PSQDF 6

RESULT 13
FAR3 CALVO
ID - FAR3 CALVO STANDARD; PRT; 9 AA.
AC P41858;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE CalliFMRamide 3.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]

RP SEQUENCE.
 RC TISSUE=Thoracic ganglion;
 RX MEDLINE=92196111; PubMed=1549595;
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A., and activity of -Phe-Met-Arg-Phe-NH2
 RT neuropeptides [designated calliFMRamides] from the blowfly
 RT Calliphora vomitoria.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330 (1992).
 CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
 CC SALIVARY GLAND OF CALLIPHORA.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC PIR; C41978; C41978.
 DR Neuropeptide; Amidation.
 KW MOD RES 9
 FT SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;
 SQ

Query Match 25.4%; Score 15; DB 1; Length 9;
 Best Local Similarity 20.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEY 5
 DB 2 PSQDF 6

RESULT 14
 FLA2_TREHY STANDARD; PRT; 9 AA.
 ID FLA2_TREHY
 AC P80159;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar filament outer layer protein fla2 (35 kDa sheath protein)
 DE (Fragment).
 GN FLA2.
 OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
 OC Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
 OX NCBI_TaxID=159;
 RN [1]

RP SEQUENCE.
 RC STRAIN=CS;
 RX MEDLINE=93139764; PubMed=1487733;
 RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
 RA van der Zeijst B.A.M., Kueters J.G.;
 RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
 RT composed of two sheath proteins and three core proteins.";
 RL J. Gen. Microbiol. 138:2697-2706 (1992).
 CC -1- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.
 CC -1- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
 CC SHEATH PROTEINS, FLA1 (44 kDa) AND FLA2 (35 kDa) AROUND A CORE
 CC THAT CONTAINS THREE PROTEINS FLA1 (37 kDa), FLA2 (34 kDa) AND
 CC FLA3 (32 kDa).
 CC -1- SUBCELLULAR LOCATION: Periplasmic flagellum.
 CC Flagella; Periplasmic.
 FT UNSURE 2
 FT UNSURE 8
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1129 MW; 855A19C68B4772D1 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 9;
 Best Local Similarity 37.5%; Pred. No. 1.3e+05;
 Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EYEYEYE 10

Db 1 ETPYMF 8

RESULT 15
 PRFL_SARBU STANDARD; PRT; 9 AA.
 ID PRFL_SARBU
 AC P83350;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neb-FMRamide 1.
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Sarcophagidae; Sarcophaga.
 OX NCBI_TaxID=7385;
 RN [1]

RP SEQUENCE, AMIDATION, AND FUNCTION.
 RC TISSUE=CNS;
 RX MEDLINE=22342733; PubMed=12438685;
 RA Meeusen T., Mertens I., Clynen E., Baggerman G., Nichols R.,
 RA Nachman R.J., Huybrechts R., De Loof A., Schoofs L.;
 RT "Identification in Drosophila melanogaster of the invertebrate G
 RT protein-coupled FMRamide receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:15363-15368 (2002).
 CC -1- FUNCTION: Has modulatory actions at skeletal neuromuscular
 CC junctions.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 1155 MW; 2D810699CAB6C5A7 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 9;
 Best Local Similarity 20.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEY 5
 DB 2 PSQDF 6

Search completed: August 5, 2003, 07:50:00
 Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:46:50 ; Search time 92 Seconds
(without alignments)
28.049 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYEEYE 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:**
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	35.6	8	13 P79940	P79940 xenopus lae
2	21	35.6	10	11 Q9QVE8	Q9QVE8 mus sp. pro
3	20	33.9	10	4 Q96QT9	Q96QT9 homo sapien
4	19	32.2	9	4 Q9UGB4	Q9UGB4 homo sapien
5	19	32.2	10	12 Q83978	Q83978 influenzavi
6	18	30.5	8	13 Q90ZV5	Q90ZV5 fulica leuc
7	16	27.1	10	5 P82384	P82384 drosophila
8	15	27.1	10	11 Q9QVK8	Q9QVK8 mus sp. mep
9	15	25.4	9	11 Q8CG13	Q8CG13 mus musculu
10	15	25.4	10	4 Q9UCP3	Q9UCP3 homo sapien
11	15	25.4	10	4 Q81ZA2	Q81ZA2 homo sapien
12	15	25.4	10	13 Q9PRY8	Q9PRY8 triakis scy
13	14	23.7	8	6 O18854	O18854 canis fami
14	14	23.7	9	3 Q9UR18	Q9UR18 sclerotium
15	14	23.7	9	4 Q9BYF9	Q9BYF9 homo sapien
16	14	23.7	9	12 Q82622	Q82622 avian infec

17	14	23.7	10	2 P83154	P83154 anabaena sp
18	14	23.7	10	4 Q8NER0	Q8NER0 homo sapien
19	14	23.7	10	4 Q96Q19	Q96Q19 homo sapien
20	14	23.7	10	4 Q8WTT4	Q8WTT4 homo sapien
21	14	23.7	10	10 P82438	P82438 nicotiana t
22	14	23.7	10	11 Q9QVJ5	Q9QVJ5 rattus sp.
23	14	23.7	10	11 Q9QVJ6	Q9QVJ6 rattus sp.
24	14	23.7	10	11 Q9QVES	Q9QVES mus sp. pro
25	13	22.0	8	2 Q9AGP4	Q9AGP4 arthrobacte
26	13	22.0	8	2 Q9R3X0	Q9R3X0 planktothri
27	13	22.0	8	2 Q94534	Q94534 mycoplasma
28	13	22.0	8	5 Q94623	Q94623 manduca sex
29	13	22.0	9	4 Q96T78	Q96T78 homo sapien
30	13	22.0	9	4 Q9UK44	Q9UK44 homo sapien
31	13	22.0	9	12 Q65711	Q65711 berne virus
32	13	22.0	9	12 Q67605	Q67605 squash leaf
33	13	22.0	9	12 Q67606	Q67606 squash leaf
34	13	22.0	9	12 Q71066	Q71066 canine dist
35	13	22.0	10	2 P96352	P96352 marinobacte
36	13	22.0	10	2 P96421	P96421 neisseria g
37	13	22.0	10	2 Q931E5	Q931E5 photobacter
38	13	22.0	10	2 P96321	P96321 escherichia
39	13	22.0	10	2 P96306	P96306 aeromonas s
40	13	22.0	10	3 Q830C2	Q830C2 encephalito
41	13	22.0	10	4 Q9UCQ8	Q9UCQ8 homo sapien
42	13	22.0	10	11 Q9QVF7	Q9QVF7 rattus sp.
43	13	22.0	10	11 Q9QVF0	Q9QVF0 mus sp. pro
44	13	22.0	10	11 Q9QVF1	Q9QVF1 mus sp. pro
45	13	22.0	10	11 Q9QVE9	Q9QVE9 mus sp. pro

ALIGNMENTS

RESULT 1
P79940 PRELIMINARY; PRT; 8 AA.
AC P79940;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE XMeisi-4 protein (fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97202105; PubMed=9049632;
RA Steelman S., Moskow J.J., Muzynski K., North C., Druck T.,
RA Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
RT "Identification of a conserved family of Meisl-related homeobox
genes";
RL Genome Res. 7:142-156(1997).
DR EMBL; U68389; AAB19199.1; --
DR TRANSFAC; T03410; --
FT NON TER 1 1
SQ SEQUENCE 8 AA; 1187 MW; 278B51F37B11F40B CRC64;
Query Match 35.6%; Score 21; DB 13; Length 8;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 EYEEY 9
Db 1 ERHEWHY 7
RESULT 2
Q9QVE8 PRELIMINARY; PRT; 10 AA.
AC Q9QVE8;

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Protamine MP2 intermediate protein PMP2/16 (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE
RX MEDLINE=92174934; PubMed=1541289;
RA Chauviere M., Martinge A., Debarle M., Sautiere P., Chevallier P.;
RT "Molecular characterization of six intermediate proteins in the
RT processing of mouse protamine p2 precursor.";
RL Eur. J. Biochem. 204:759-765(1992).
FT NON_TER 1
FT NON_TER 10
SQ SEQUENCE 10 AA; 1224 MW; D4050B040B11EAB6 CRC64;

Query Match 35.6%; Score 21; DB 11; Length 10;
Best Local Similarity 37.5%; Pred. No. 2.9e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEYEYE 8
Db 1 PGQDHERE 8

RESULT 3
Q96QT9 PRELIMINARY; PRT; 10 AA.
AC Q96QT9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Protein tyrosine phosphatase 1B (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21326081; PubMed=11316810;
RA Fukuda T., Tonks N.K.;
RT "The Reciprocal Role of Egr-1 and Sp Family Proteins in Regulation of
RT the PTP1B Promoter in Response to the p210 Bcr-Abl Oncoprotein-
RT tyrosine Kinase.";
RL J. Biol. Chem. 276:25512-25519(2001).
DR EMBL; AY029236; AAK31734.1; -.
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1314 MW; AC08E219CB133B16 CRC64;

Query Match 33.9%; Score 20; DB 4; Length 10;
Best Local Similarity 57.1%; Pred. No. 4.2e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 EYEYEYE 10
Db 2 EMEKEFE 8

RESULT 4
Q9UGE4 PRELIMINARY; PRT; 9 AA.
AC Q9UGE4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE DJ34105.1 (Remainder of gene in sequence AL023513) (Fragment).
GN SEZ6L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.; (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL078460; CAB51751.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 9 AA; 1125 MW; 8154A2CB05B0411 CRC64;

Query Match 32.2%; Score 19; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EYEYE 8
Db 1 ETEYE 6

RESULT 5
Q83978 PRELIMINARY; PRT; 10 AA.
AC Q83978;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Influenza A/udorn/72 (H3n2), nucleoprotein (Seg 5), 3' cdna
DE (Fragment).
OS Influenzavirus A.
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses.
OX NCBI_TaxID=197911;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112211; PubMed=6296449;
RA Lin B.-C., Lai C.-J.;
RT "the influenza virus nucleoprotein synthesized from cloned dna in a
RT simian virus 40 vector is detected in the nucleus.";
RL J. Virol. 45:434-438(1983).
DR EMBL; J02171; AAA43469.1; -.
KW Nucleocapsid.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 10 AA; 1173 MW; 8787655B1B1DD44A CRC64;

Query Match 32.2%; Score 19; DB 12; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.1e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EYE 6
Db 6 EYD 9

RESULT 6
Q90ZV5 PRELIMINARY; PRT; 8 AA.
AC Q90ZV5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Adenylate kinase (Fragment).
OS Fulica leucoptera.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Fulica.
OX NCBI_TaxID=156758;
RN [1]
RP SEQUENCE FROM N.A.
RA Shapiro L.H., Dumbacher J.P.;
RT "Adenylate kinase intron 5: a new nuclear locus for avian
RT systematics.";
RL Auk 118:248-255(2001).
DR EMBL; AF307898; AAK43537.1; -.
KW Kinase.
FT NON_TER 1
FT NON_TER 1

```

FT NON TER 8 8
SQ SEQUENCE 8 AA; 994 MW; 96333B19CB1B1866 CRC64;
Query Match 30.5%; Score 18; DB 13; Length 8;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 BEYE 6
Db 3 EEPE 6

RESULT 7
ID P2384 PRELIMINARY; PRT; 10 AA.
AC P2384;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Larval cuticle Lcp9 protein (Fragment).
GN LCP9.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE.
RC STRAIN=OREGON-R; TISSUE=Larva;
RA Chihara C.J.;
RT "Third instar cuticle proteins.";
RL DIS 83:0-0(2000).
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF THE LARVA OF FRUIT FLY.
DR FlyBase; FBgn025578; Lcp9.
DR InterPro; IPR000618; Insect cuticle.
DR PROSITE; PS00233; CUTICLE; PARTIAL.
KW Cuticle; Structural protein.
FT NON TER 10
SQ SEQUENCE 10 AA; 1105 MW; 7EF84522D2CAADB CRC64;

Query Match 27.1%; Score 16; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.9e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NEE 4
Db 1 NEE 3

RESULT 8
ID Q9QVK8 PRELIMINARY; PRT; 10 AA.
AC Q9QVK8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MEPRIN=METALLOENDOPEPTIDASE (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=91363409; PubMed=1888759;
RA Flannery A.V., Macadam G.C., Beynon R.J.;
RT "Immunological characterisation of different meprin species in mice.";
RL Biochim. Biophys. Acta 1079:119-122(1991).
FT NON TER 1
FT NON TER 10
SQ SEQUENCE 10 AA; 1163 MW; DD6436144731B2C9 CRC64;

Query Match 27.1%; Score 16; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.9e+04;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 NEE 4
Db 7 NEE 9

RESULT 9
ID Q8CG13 PRELIMINARY; PRT; 9 AA.
AC Q8CG13;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glutamate receptor ionotropic N-methyl D-aspartate-like 1A (Fragment).
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Wyden K.S., Mohan Raj B.K., Sciorra L.J., Roginski R.S.;
RT "The mouse orthologue of the human ionotropic glutamate receptor-like gene (GRIN1A) maps to mouse chromosome 9.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF462417; AAO15648.1;
DR EMBL; AF462416; AAO15648.1; JOINED.
KW Receptor.
FT NON TER 1
FT NON TER 9
SQ SEQUENCE 9 AA; 1091 MW; 6A91233EB059C33B CRC64;

Query Match 25.4%; Score 15; DB 11; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NEEY 5
Db 1 NEKF 4

RESULT 10
ID Q9UCP3 PRELIMINARY; PRT; 10 AA.
AC Q9UCP3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE cGMP-inhibited LOW K(M) CAMP phosphodiesterase PEAK 37, CGI-PDE (Fragment).
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=92283180; PubMed=1317779;
RA LeBon T.R., Kasuya J., Paxton R.J., Belfrage P., Hockman S., Manganiello V.C., Fujita Yamaguchi Y.;
RT "Purification and characterization of guanosine 3',5'-monophosphate-inhibited low K(m) adenosine 3',5'-monophosphate phosphodiesterase from human placental cytosolic fractions.";
RL Endocrinology 130:3265-3274(1992).
FT NON TER 1
FT NON TER 10
SQ SEQUENCE 10 AA; 1272 MW; C80C440B5449C046 CRC64;

Query Match 25.4%; Score 15; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 2.7e+04;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 YEYE 8
Db 5 FYYE 8

RESULT 11

ID Q812A2 PRELIMINARY; PRT; 10 AA.
AC Q812A2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cardiac troponin T (Fragment).
GN TNNT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Bahl A., Kubo T., Steffensen U., Steffensen M., McKenna W.J.,
Mogensen J.;
RT "Homo sapiens Troponin T (TNNT2) introns 6 and 7 and exon 7."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY160215; AAN71650.1; -.
FT NON_TER 1
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1099 MW; AB24BC6325BB1B16 CRC64;

Query Match 25.4%; Score 15; DB 4; Length 10;
Best Local Similarity 75.0%; Pred. No. 2.7e+04;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEE 4
Db 2 PMEE 5

RESULT 12

ID Q9PRY8 PRELIMINARY; PRT; 10 AA.
AC Q9PRY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Angiotensin I.
OS Triakis scyllium (Leopard shark) (Triakis scyllia).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC Triakis.
OX NCBI_TaxID=30494;
RN [1]
RP SEQUENCE.
RX MEDLINE=9411412; PubMed=8308464;
RA Takei Y., Hasegawa Y., Watanabe T.X., Nakajima K., Hazon N.;
RT "A novel angiotensin I isolated from an elasmobranch fish."
RL J. Endocrinol. 139:281-285(1993).
SQ SEQUENCE 10 AA; 1284 MW; 20F02FD761E04B47 CRC64;

Query Match 25.4%; Score 15; DB 13; Length 10;
Best Local Similarity 22.2%; Pred. No. 2.7e+04;
Matches 2; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NEEYEYE 10
Db 1 NRPYHPFQ 9

RESULT 13

O18854
ID O18854 PRELIMINARY; PRT; 8 AA.
AC O18854;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CD19 antigen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu P.-C., Shibuya H., Katz M.L., Johnson G.S.;
RT "A BsmFI PCR/RELP in the canine CD19 antigen gene."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024717; AAB81967.1; -.
FT NON_TER 1
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 832 MW; 6735A1ADDB1325A7 CRC64;

Query Match 23.7%; Score 14; DB 6; Length 8;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEE 4
Db 1 PSKE 4

RESULT 14

ID Q9UR18 PRELIMINARY; PRT; 9 AA.
AC Q9UR18;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Translation elongation factor 1-alpha (Fragment).
GN EF1-ALPHA.
OS Sclerotium cepivorum.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Sclerotium.
OX NCBI_TaxID=38492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DE-2, and DE-1;
RA Couch B.C., Kohn L.M.;
RT "Clonal spread of Sclerotium cepivorum in onion production with
evidence of past recombination events."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163671; AAD46912.1; -.
DR EMBL; AF163670; AAD46911.1; -.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 9
FT NON_TER 9
SQ SEQUENCE 9 AA; 1178 MW; 619C19C33B1041B4 CRC64;

Query Match 23.7%; Score 14; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EEEY 6
Db 5 EKPE 8

RESULT 15

ID Q9BYF9 PRELIMINARY; PRT; 9 AA.
AC Q9BYF9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytokeratin 19 (Fragment).
GN K19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21539745; PubMed=11682035;
RA Kagawa M., Kaneko S., Ohno H., Inamura K., Kobayashi K.;
RT "Cloning and characterization of the 5'-flanking region of human
cytokerin 19 gene in human cholangiocarcinoma cell line.";
RL J. Hepatol. 35:504-511 (2001).
DR EMBL; AB045973; BAB40770.1; -.
KW Keratin.
FT NON TER
SQ SEQUENCE 9 AA; 1122 MW; 9E9FC41B45AB45A1 CRC64;

Query Match 23.7%; Score 14; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 YYY 7
Db 4 YSY 6

Search completed: August 5, 2003, 07:51:41
Job time : 94 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:39:59 ; Search time 81 Seconds
(without alignments)
19.596 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYEEYE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	10	23	AAE20072
2	56	94.9	10	23	AAE19949
3	56	94.9	10	23	AAE19957
4	56	94.9	10	23	AAE19958
5	56	94.9	10	23	AAE19959
6	56	94.9	10	23	AAE19960
7	56	94.9	10	23	AAE19961
8	56	94.9	10	23	AAE19962
9	56	94.9	10	23	AAE19963
					Complement C1s inh
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo
					Complement C1s exo

10	52	88.1	10	23	AAE19950	Complement C1s exo
11	52	88.1	10	23	AAE19951	Complement C1s exo
12	52	88.1	10	23	AAE19952	Complement C1s exo
13	52	88.1	10	23	AAE19964	Complement C1s exo
14	52	88.1	10	23	AAE19965	Complement C1s exo
15	52	88.1	10	23	AAE19966	Complement C1s exo
16	52	88.1	10	23	AAE19967	Complement C1s exo
17	52	88.1	10	23	AAE19968	Complement C1s exo
18	52	88.1	10	23	AAE19969	Complement C1s exo
19	52	88.1	10	23	AAE19972	Complement C1s exo
20	52	88.1	10	23	AAE19974	Complement C1s exo
21	48	81.4	10	23	AAE19953	Complement C1s exo
22	48	81.4	10	23	AAE19954	Complement C1s exo
23	48	81.4	10	23	AAE19955	Complement C1s exo
24	48	81.4	10	23	AAE19970	Complement C1s exo
25	48	81.4	10	23	AAE19971	Complement C1s exo
26	48	81.4	10	23	AAE19973	Complement C1s exo
27	44	74.6	10	23	AAE19956	Complement C1s exo
28	35.5	60.2	10	23	AAE20066	Complement C1s exo
29	33	55.9	10	22	AAE20066	Human complement
30	29	49.2	8	15	AAE57388	Peptide for treati
31	29	49.2	9	20	AAW97641	Frazzled protein W
32	29	49.2	9	22	AAE91235	Cholecystokinin (C
33	29	49.2	10	15	AAE57407	Peptide for treati
34	29	49.2	10	15	AAE57394	Peptide for treati
35	29	49.2	10	20	AAE40997	Ara h 3 allergen I
36	29	49.2	10	21	AAE27494	Ara h 3 peptide #3
37	29	49.2	10	24	ABU52491	Peanut Ara h3 IGE
38	28	47.5	9	14	AAE30827	Sequence of peptid
39	28	47.5	9	18	AAW54327	Bradykinin analogo
40	28	47.5	9	19	AAW77424	Thrombin receptor
41	28	47.5	10	16	AAE74936	H-CDR-3 of anti-id
42	28	47.5	10	22	AAE43101	Mycoplasma genital
43	26	44.1	9	15	AAE67132	Factor Xa derived
44	26	44.1	9	22	AAU06371	Human Leukocyte An
45	26	44.1	9	22	AAE65312	Anti-IL-18 antibod

ALIGNMENTS

RESULT 1
AAE20072
ID AAE20072 standard; peptide; 10 AA.
XX
AC AAE20072;
XX
DT 18-JUN-2002 (first entry)
XX
Complement C1s inhibitor peptide #1.

DE
DE
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejectin; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
restenosis; myasthenia gravis.
XX
XX Unidentified.
XX
XX WO200198365-A2.
XX
XX 27-DEC-2001.
XX
XX 18-JUN-2001; 2001WO-US19405.
XX
XX 21-JUN-2000; 2000US-212998P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX West RR, Sheppard PO, Fox BA;
XX
XX WPI; 2002-241177/29.
XX

PT New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
PT sepsis -
XX
PS Claim 4; Page 49; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s inhibitor peptide.
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 59; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0012; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;
Qy 1 PNEEYEEYE 10
Db 1 PNEEYEEYE 10
|||||:|
|||||:|
RESULT 2
AAE19949
ID AAE19949 standard; peptide; 10 AA.
XX
AC AAE19949;
XX
DT 18-JUN-2002 (first entry)
XX
DE Complement C1s exosite binding peptide #1.
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
OS Unidentified.
XX
FN WO200198365-A2.
XX
PD 27-DEC-2001.
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
PR 21-JUN-2000; 2000US-212998P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI West RR, Sheppard PO, Fox BA;
XX
DR WPI; 2002-241177/29.
XX
CC New complement C1s inhibitor polypeptides for treating diseases in
CC which complement activation has been shown to occur, e.g. adult
CC respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
CC sepsis -
XX
PS Disclosure; Page 15; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
SQ Sequence 10 AA;
Query Match 94.9%; Score 56; DB 23; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0037; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 1;
Qy 1 PNEEYEEYE 10
Db 1 PNEEYEEYDYE 10
|||||:|
|||||:|
RESULT 3
AAE19957
ID AAE19957 standard; peptide; 10 AA.
XX
AC AAE19957;
XX
DT 18-JUN-2002 (first entry)
XX
DE Complement C1s exosite binding peptide #9.
XX
KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.
XX
OS Unidentified.
XX
FN Key Location/Qualifiers
FT Modified-site 5 /note= "Sulphated Tyr"
FT
XX WO200198365-A2.
XX
PD 27-DEC-2001.
XX
PF 18-JUN-2001; 2001WO-US19405.
XX
PR 21-JUN-2000; 2000US-212998P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI West RR, Sheppard PO, Fox BA;
XX
DR WPI; 2002-241177/29.
XX
CC New complement C1s inhibitor polypeptides for treating diseases in
CC which complement activation has been shown to occur, e.g. adult
CC respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
CC sepsis -
XX
PS Disclosure; Page 15; 99pp; English.
XX
CC The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.
XX
SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYEDYE 10

RESULT 4

AAE19958
 ID AAE19958 standard; peptide; 10 AA.

XX
 AC AAE19958;

XX 18-JUN-2002 (first entry)

XX Complement C1s exosite binding peptide #10.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute relectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

XX Unidentified.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Sulphated Tyr"

XX WO200198365-A2.

XX 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis -

XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute relectin, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

XX Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYEDYE 10

RESULT 5

AAE19959
 ID AAE19959 standard; peptide; 10 AA.

XX
 AC AAE19959;

XX 18-JUN-2002 (first entry)

XX Complement C1s exosite binding peptide #11.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute relectin; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.

XX Unidentified.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Sulphated Tyr"

XX WO200198365-A2.

XX 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
 PT sepsis -

XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute relectin, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

XX Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0037;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
 |||||:|
 Db 1 PNEEYEDYE 10

RESULT 6

AAE19960
 ID AAE19960 standard; peptide; 10 AA.

XX
 AC AAE19960;

XX 18-JUN-2002 (first entry)

XX DE Complement C1s exosite binding peptide #12.

XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute resection; rheumatoid arthritis; burn; wound healing; asthma;

XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

XX KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Modified-site 5

FT Modified-site 7 /note= "Sulphated Tyr"

FT Modified-site 7 /note= "Sulphated Tyr"

XX WO200198365-A2.

XX 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (Zymo) ZYMOGENETICS INC.

PI West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in

PT which complement activation has been shown to occur, e.g. adult

PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or

PT sepsis -

XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The

CC inhibitory peptides are useful as therapeutic agents, as preservatives

CC in blood samples, and in affinity purification procedures to isolate C1s.

CC Molecules that inhibit complement may be used for treating diseases

CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial

CC infarct, stroke), hyperacute resection, sepsis, burns, wound healing,

CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present

CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;

Best Local Similarity 90.0%; Pred. No. 0.0037;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PNEYEY EYE 10

DB 1 PNEYEYDY E 10

RESULT 7

AAE19961

XX AAE19961 standard; peptide; 10 AA.

XX AAE19961;

XX 18-JUN-2002 (first entry)

XX Complement C1s exosite binding peptide #12.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute resection; rheumatoid arthritis; burn; wound healing; asthma;

KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

XX restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Modified-site 5

FT Modified-site 9 /note= "Sulphated Tyr"

FT Modified-site 9 /note= "Sulphated Tyr"

XX WO200198365-A2.

XX 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (Zymo) ZYMOGENETICS INC.

PI West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in

PT which complement activation has been shown to occur, e.g. adult

PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or

PT sepsis -

XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The

CC inhibitory peptides are useful as therapeutic agents, as preservatives

CC in blood samples, and in affinity purification procedures to isolate C1s.

CC Molecules that inhibit complement may be used for treating diseases

CC in which complement activation has been shown to occur, e.g., adult

CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial

CC infarct, stroke), hyperacute resection, sepsis, burns, wound healing,

CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present

CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 94.9%; Score 56; DB 23; Length 10;

Best Local Similarity 90.0%; Pred. No. 0.0037;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PNEYEY EYE 10

DB 1 PNEYEYDY E 10

RESULT 8

AAE19962

XX AAE19962 standard; peptide; 10 AA.

XX AAE19962;

XX 18-JUN-2002 (first entry)

XX Complement C1s exosite binding peptide #12.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;

XX KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;

XX KW hyperacute resection; rheumatoid arthritis; burn; wound healing; asthma;

XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;

XX restenosis; myasthenia gravis.

XX OS Unidentified.

XX FH Key Location/Qualifiers


```

XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI West RR, Sheppard PO, Fox BA;
XX DR WPI; 2002-241177/29.
XX XX
XX PT New complement C1s inhibitor polypeptides for treating diseases in
XX PT which complement activation has been shown to occur, e.g. adult
XX PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
XX PT sepsis -
XX PS Disclosure; Page 15; 99pp; English.
XX XX
XX CC The invention relates to a polypeptide that inhibits complement C1s. The
XX CC inhibitory peptides are useful as therapeutic agents, as preservatives
XX CC in blood samples, and in affinity purification procedures to isolate C1s.
XX CC Molecules that inhibit complement may be used for treating diseases
XX CC in which complement activation has been shown to occur, e.g., adult
XX CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
XX CC infarct, stroke), hyperacute resection, sepsis, burns, wound healing,
XX CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
XX CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
XX CC sequence is complement C1s exosite binding peptide.
XX SQ Sequence 10 AA;
Query Match 88.1%; Score 52; DB 23; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.017;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PNEEYVEYE 10
Db 1 PNEEYFDYE 10
|||||:|
|||||:|

RESULT 11
AAE19951
ID AAE19951 standard; peptide; 10 AA.
AC AAE19951;
XX XX
XX DT 18-JUN-2002 (first entry)
XX DE Complement C1s exosite binding peptide #3.
XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
XX KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
XX KW hyperacute resection; rheumatoid arthritis; burn; wound healing; asthma;
XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
XX KW restenosis; myasthenia gravis.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT Modified-site 7 /note= "Sulphated Phe"
XX FT FT
XX PN WO200198365-A2.
XX XX
XX PD 27-DEC-2001.
XX XX
XX PF 18-JUN-2001; 2001WO-US19405.
XX XX
XX PR 21-JUN-2000; 2000US-212998P.
XX XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI West RR, Sheppard PO, Fox BA;
XX XX
XX DR WPI; 2002-241177/29.
XX XX
XX PT New complement C1s inhibitor polypeptides for treating diseases in
XX PT which complement activation has been shown to occur, e.g. adult
XX PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
XX PT sepsis -
XX PS Disclosure; Page 15; 99pp; English.
XX XX
XX CC The invention relates to a polypeptide that inhibits complement C1s. The
XX CC inhibitory peptides are useful as therapeutic agents, as preservatives
XX CC in blood samples, and in affinity purification procedures to isolate C1s.
XX CC Molecules that inhibit complement may be used for treating diseases
XX CC in which complement activation has been shown to occur, e.g., adult
XX CC respiratory distress syndrome, ischemia-reperfusion injury (myocardial
XX CC infarct, stroke), hyperacute resection, sepsis, burns, wound healing,
XX CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
XX CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
XX CC sequence is complement C1s exosite binding peptide.
XX SQ Sequence 10 AA;
Query Match 88.1%; Score 52; DB 23; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.017;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PNEEYVEYE 10
Db 1 PNEEYFDYE 10
|||||:|
|||||:|

RESULT 12
AAE19952
ID AAE19952 standard; peptide; 10 AA.
AC AAE19952;
XX XX
XX DT 18-JUN-2002 (first entry)
XX DE Complement C1s exosite binding peptide #4.
XX KW Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
XX KW preservative; ischemia-reperfusion injury; myocardial infarct; sepsis;
XX KW hyperacute resection; rheumatoid arthritis; burn; wound healing; asthma;
XX KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
XX KW restenosis; myasthenia gravis.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT Modified-site 9 /note= "Sulphated Phe"
XX FT FT
XX PN WO200198365-A2.
XX XX
XX PD 27-DEC-2001.
XX XX
XX PF 18-JUN-2001; 2001WO-US19405.
XX XX
XX PR 21-JUN-2000; 2000US-212998P.
XX XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI West RR, Sheppard PO, Fox BA;
XX XX
XX DR WPI; 2002-241177/29.
XX XX
XX PT New complement C1s inhibitor polypeptides for treating diseases in
XX PT which complement activation has been shown to occur, e.g. adult
XX PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
XX PT sepsis -
XX PS Disclosure; Page 15; 99pp; English.
XX XX
XX CC The invention relates to a polypeptide that inhibits complement C1s. The
XX CC inhibitory peptides are useful as therapeutic agents, as preservatives

```

CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.
 XX
 SQ Sequence 10 AA;

Query Match 88.1%; Score 52; DB 23; Length 10;
 Best Local Similarity 80.0%; Pred. NO. 0.017; Indels 0; Gaps 0;
 Matches 8; Conservative 2; Mismatches 0;

Qy 1 PNEEYEEYE 10
 |||||:|:
 Db 1 PNEEYEDFE 10

RESULT 13
 AAE19964
 ID AAE19964 standard; peptide; 10 AA.
 XX
 AC AAE19964;

18-JUN-2002 (first entry)

Complement C1s exosite binding peptide #16.

CC Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.

Key Location/Qualifiers
 Modified-site 5 /note= "Sulphated Phe"
 Modified-site 7 /note= "Sulphated Tyr"

WO200198365-A2.

27-DEC-2001.

18-JUN-2001; 2001WO-US19405.

21-JUN-2000; 2000US-212998P.

(ZYMO) ZYMOGENETICS INC.

West RR, Sheppard PO, Fox BA;

WPI; 2002-241177/29.

New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
 PT sepsis -

Disclosure; Page 15; 99pp; English.

CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,

CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

SQ Sequence 10 AA;

Query Match 88.1%; Score 52; DB 23; Length 10;
 Best Local Similarity 80.0%; Pred. NO. 0.017; Indels 0; Gaps 0;
 Matches 8; Conservative 2; Mismatches 0;

Qy 1 PNEEYEEYE 10
 |||||:|:
 Db 1 PNEEYEDYE 10

RESULT 14
 AAE19965
 ID AAE19965 standard; peptide; 10 AA.
 XX
 AC AAE19965;

18-JUN-2002 (first entry)

Complement C1s exosite binding peptide #17.

CC Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
 KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
 KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
 KW systemic lupus erythematosus; SLE; haemolytic anaemia; serum sickness;
 KW restenosis; myasthenia gravis.
 XX
 OS Unidentified.

Key Location/Qualifiers
 Modified-site 5 /note= "Sulphated Tyr"
 Modified-site 7 /note= "Sulphated Phe"

WO200198365-A2.

27-DEC-2001.

18-JUN-2001; 2001WO-US19405.

21-JUN-2000; 2000US-212998P.

(ZYMO) ZYMOGENETICS INC.

West RR, Sheppard PO, Fox BA;

WPI; 2002-241177/29.

New complement C1s inhibitor polypeptides for treating diseases in
 PT which complement activation has been shown to occur, e.g. adult
 PT respiratory distress syndrome, ischaemia-reperfusion injury, asthma, or
 PT sepsis -

Disclosure; Page 15; 99pp; English.

CC The invention relates to a polypeptide that inhibits complement C1s. The
 CC inhibitory peptides are useful as therapeutic agents, as preservatives
 CC in blood samples, and in affinity purification procedures to isolate C1s.
 CC Molecules that inhibit complement may be used for treating diseases
 CC in which complement activation has been shown to occur, e.g., adult
 CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
 CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
 CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
 CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
 CC sequence is complement C1s exosite binding peptide.

SQ Sequence 10 AA;

Query Match 88.1%; Score 52; DB 23; Length 10;

Best Local Similarity 80.0%; Pred. No. 0.017; Mismatches 2; Indels 0; Gaps 0;

Search completed: August 5, 2003, 07:49:29
Job time : 81 secs

QY 1 PNEYEYEE 10
DB 1 PNEYEYDYE 10

RESULT 15

AAE19966
ID AAE19966 standard; peptide; 10 AA.

XX AC AAE19966;

XX 18-JUN-2002 (first entry)

XX Complement C1s exosite binding peptide #18.

XX Complement C1s; therapeutic; adult respiratory distress syndrome; stroke;
KW preservative; ischaemia-reperfusion injury; myocardial infarct; sepsis;
KW hyperacute rejection; rheumatoid arthritis; burn; wound healing; asthma;
KW systemic lupus erythematosus; SL8; haemolytic anaemia; serum sickness;
KW restenosis; myasthenia gravis.

XX OS Unidentified.

XX Key Location/Qualifiers

FT Modified-site 5 /note= "Sulphated Phe"

FT Modified-site 9 /note= "Sulphated Tyr"

XX WO200198365-A2.

XX 27-DEC-2001.

XX 18-JUN-2001; 2001WO-US19405.

XX 21-JUN-2000; 2000US-212998P.

XX (ZYMO) ZYMOGENETICS INC.

XX West RR, Sheppard PO, Fox BA;

XX WPI; 2002-241177/29.

XX New complement C1s inhibitor polypeptides for treating diseases in
PT which complement activation has been shown to occur, e.g. adult
PT respiratory distress syndrome, ischemia-reperfusion injury, asthma, or
PT sepsis

XX Disclosure; Page 15; 99pp; English.

XX The invention relates to a polypeptide that inhibits complement C1s. The
CC inhibitory peptides are useful as therapeutic agents, as preservatives
CC in blood samples, and in affinity purification procedures to isolate C1s.
CC Molecules that inhibit complement may be used for treating diseases
CC in which complement activation has been shown to occur, e.g., adult
CC respiratory distress syndrome, ischaemia-reperfusion injury (myocardial
CC infarct, stroke), hyperacute rejection, sepsis, burns, wound healing,
CC asthma, restenosis, rheumatoid arthritis, systemic lupus erythematosus,
CC serum sickness, haemolytic anaemias, or myasthenia gravis. The present
CC sequence is complement C1s exosite binding peptide.

XX SQ Sequence 10 AA;

Query Match 88.1%; Score 52; DB 23; Length 10;

Best Local Similarity 80.0%; Pred. No. 0.017; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEYEYEE 10
DB 1 PNEYEYDYE 10

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 07:51:45 ; Search time 49 Seconds
(without alignments)

24.237 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59
Sequence: 1 PNEYEYEYE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 58525

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	10	10	US-09-883-727A-125
2	56	94.9	10	10	US-09-883-727A-2
3	48	81.4	10	10	US-09-883-727A-3
4	48	81.4	10	10	US-09-883-727A-4
5	48	81.4	10	10	US-09-883-727A-5
6	48	81.4	10	10	US-09-883-727A-10
7	48	81.4	10	10	US-09-883-727A-11
8	48	81.4	10	10	US-09-883-727A-12
9	40	67.8	10	10	US-09-883-727A-6
10	40	67.8	10	10	US-09-883-727A-7
11	40	67.8	10	10	US-09-883-727A-8
12	40	67.8	10	10	US-09-883-727A-13
13	40	67.8	10	10	US-09-883-727A-14
14	40	67.8	10	10	US-09-883-727A-15
15	40	67.8	10	10	US-09-883-727A-17

16	40	67.8	10	10	US-09-883-727A-18
17	40	67.8	10	10	US-09-883-727A-19
18	40	67.8	10	10	US-09-883-727A-20
19	40	67.8	10	10	US-09-883-727A-21
20	40	67.8	10	10	US-09-883-727A-22
21	35.5	60.2	10	10	US-09-883-727A-119
22	33	55.9	10	11	US-09-572-404B-892
23	32	54.2	10	10	US-09-883-727A-9
24	32	54.2	10	10	US-09-883-727A-16
25	32	54.2	10	10	US-09-883-727A-23
26	32	54.2	10	10	US-09-883-727A-24
27	32	54.2	10	10	US-09-883-727A-25
28	32	54.2	10	10	US-09-883-727A-26
29	32	54.2	10	10	US-09-883-727A-27
30	25	42.4	7	15	US-10-305-268-7
31	25	42.4	9	10	US-09-252-150-76
32	25	42.4	10	15	US-10-003-983B-16
33	24	40.7	10	15	US-10-128-711-8
34	23.5	39.8	10	9	US-09-834-765-93
35	23	39.0	9	15	US-10-094-699-62
36	23	39.0	9	15	US-10-219-850-20
37	23	39.0	10	15	US-10-094-699-63
38	22	37.3	5	10	US-09-234-395-182
39	22	37.3	5	10	US-09-305-928-182
40	22	37.3	5	10	US-09-264-516A-206
41	22	37.3	5	14	US-10-020-436A-6
42	22	37.3	5	15	US-10-006-869-796
43	22	37.3	5	15	US-10-006-869-1224
44	22	37.3	5	15	US-10-006-869-2528
45	22	37.3	5	15	US-10-141-357-206

ALIGNMENTS

RESULT 1
US-09-883-727A-125
; Sequence 125, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Complement C1s inhibitor
US-09-883-727A-125

Query Match 100.0%; Score 59; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEYEYEYE 10
|||||
DB 1 PNEYEYEYE 10

RESULT 2
US-09-883-727A-2
; Sequence 2, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.

```
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
US-09-883-727A-2

Query Match      94.9%; Score 56; DB 10; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0046;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
Db 1 PNEEYDYDYE 10

RESULT 3
US-09-883-727A-3
; Sequence 3, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-3

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
Db 1 PNEEYDYDYE 10

RESULT 4
US-09-883-727A-4
; Sequence 4, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-4

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
Db 1 PNEEYXDYE 10

RESULT 5
US-09-883-727A-5
; Sequence 5, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-5

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
Db 1 PNEEYXDYE 10

RESULT 6
US-09-883-727A-10
; Sequence 10, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
```



```
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-12

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
   |||||:|
Db 1 PNEEYEDYE 10

RESULT 7
US-09-883-727A-11
; Sequence 11, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-11

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
   |||||:|
Db 1 PNEEYEDYE 10

RESULT 8
US-09-883-727A-12
; Sequence 12, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 10
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-125

Query Match      81.4%; Score 48; DB 10; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.083;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
   |||||:|
Db 1 PNEEYEDYE 10

RESULT 9
US-09-883-727A-6
; Sequence 6, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-6

Query Match      67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
   |||||:|
Db 1 PNEEYEDYE 10

RESULT 10
US-09-883-727A-7
; Sequence 7, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-7

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|||||:|
Db 1 PNEEXEYDXE 10

RESULT 11

US-09-883-727A-8
; Sequence 8, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = Phe-(p-CH2)SO3H
US-09-883-727A-8

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|||||:|
Db 1 PNEEXEYDXE 10

RESULT 12

US-09-883-727A-13
; Sequence 13, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 13
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = sulfated tyrosine
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-13

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|||||:|
Db 1 PNEEXEYDXE 10

RESULT 13

US-09-883-727A-14
; Sequence 14, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = sulfated tyrosine
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-14

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEEYE 10
|||||:|
Db 1 PNEEXEYDXE 10

RESULT 14

US-09-883-727A-15
; Sequence 15, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A

; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = sulfated tyrosine
; NAME/KEY: MUTAGEN
; LOCATION: (9)...(9)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-15

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
|||||:
Db 1 PNEEYEXDE 10

RESULT 15

US-09-883-727A-17
; Sequence 17, Application US/09883727A
; Patent No. US20020102256A1
; GENERAL INFORMATION:
; APPLICANT: West, Robert R.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: Peptide and Polypeptide Inhibitors of
; TITLE OF INVENTION: Complement C1s
; FILE REFERENCE: 00-33
; CURRENT APPLICATION NUMBER: US/09/883,727A
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C1s exosite binding moiety
; NAME/KEY: MUTAGEN
; LOCATION: (5)...(5)
; OTHER INFORMATION: Xaa = Phe- (p-CH2)SO3H
; NAME/KEY: MUTAGEN
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = sulfated tyrosine
US-09-883-727A-17

Query Match 67.8%; Score 40; DB 10; Length 10;
Best Local Similarity 70.0%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PNEEYEEYE 10
|||||:
Db 1 PNEEYEXDE 10

Search completed: August 5, 2003, 08:00:34
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw_model

Run on: August 5, 2003, 07:48:05 ; Search time 29 Seconds
(without alignments)
14.590 Million cell updates/sec

Title: US-09-883-727A-125

Perfect score: 59

Sequence: 1 PNEYEYVEYE 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	49.2	9	3	US-08-893-654B-17
2	28	47.5	9	1	US-08-447-010-9
3	28	47.5	9	3	US-08-676-242-3
4	28	47.5	9	3	US-09-402-732-3
5	28	47.5	10	1	US-08-318-970B-8
6	27	45.8	9	2	US-08-318-856A-20
7	26	44.1	9	1	US-07-841-997A-6
8	26	44.1	9	1	US-08-290-301-6
9	26	44.1	9	2	US-08-934-222-125
10	26	44.1	9	2	US-08-933-402-125
11	26	44.1	9	2	US-09-207-621-125
12	26	44.1	9	2	US-08-532-818-125
13	26	44.1	9	3	US-09-231-797-125
14	26	44.1	9	3	US-08-934-224-125
15	26	44.1	9	3	US-08-933-843-125
16	26	44.1	9	3	US-08-934-223-125
17	26	44.1	9	3	US-09-413-492-125
18	26	44.1	9	4	US-09-013-598-6
19	26	44.1	10	1	US-08-365-394-1
20	26	44.1	10	3	US-08-159-339A-935
21	26	44.1	10	5	PCT-US91-01543-1
22	24	40.7	9	2	US-08-318-856A-28
23	24	40.7	10	1	US-08-520-977A-5
24	24	40.7	10	4	US-08-197-484-8
25	24	40.7	10	5	PCT-US95-02121-8
26	24	40.7	10	5	PCT-US95-04975-18
27	23	39.0	7	2	US-08-408-095-3

28	23	39.0	7	4	US-09-170-769A-12	Sequence 12, Appl
29	23	39.0	8	3	US-09-222-373-37	Sequence 37, Appl
30	23	39.0	8	3	US-09-222-373-42	Sequence 42, Appl
31	23	39.0	8	3	US-09-001-511-37	Sequence 37, Appl
32	23	39.0	8	3	US-09-001-511-42	Sequence 42, Appl
33	23	39.0	8	4	US-09-510-616-37	Sequence 37, Appl
34	23	39.0	8	4	US-09-510-616-42	Sequence 42, Appl
35	23	39.0	9	2	US-08-466-337A-3	Sequence 3, Appl
36	23	39.0	9	2	US-08-475-359-3	Sequence 3, Appl
37	23	39.0	9	3	US-08-465-887A-3	Sequence 3, Appl
38	23	39.0	9	3	US-08-159-339A-691	Sequence 691, App
39	23	39.0	9	4	US-08-634-332A-78	Sequence 78, Appl
40	23	39.0	10	3	US-08-159-339A-657	Sequence 657, App
41	23	39.0	10	3	US-09-064-713-2	Sequence 2, Appl
42	22	37.3	5	4	US-09-187-859-796	Sequence 796, App
43	22	37.3	5	4	US-09-187-859-1224	Sequence 1224, Ap
44	22	37.3	5	4	US-09-187-859-2528	Sequence 2528, Ap
45	22	37.3	5	4	US-09-305-927-206	Sequence 206, App

ALIGNMENTS

RESULT 1
US-08-893-654B-17
; Sequence 17, Application US/08893654B
; Patent No. 6165748
; GENERAL INFORMATION:
; APPLICANT: RACIE, LISA, ET ALIA
; TITLE OF INVENTION: Frazzled NUCLEOTIDE SEQUENCES,
; TITLE OF INVENTION: EXPRESSION PRODUCTS, COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGE PARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,654B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEINERT, M.C.
; REGISTRATION NUMBER: 31,544
; REFERENCE/DOCKET NUMBER: GI 5279
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617.498.8574
; TELEFAX: 617.876.5851
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-893-654B-17

Query Match 49.2%; Score 29; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 3 EYEYVEYE 10
Db 1 KEYQYAYK 8

RESULT 2
US-08-447-010-9
; Sequence 9, Application US/08447010
; Patent No. 5770718
; GENERAL INFORMATION:
; APPLICANT: MOFFATT, BARBARA
; TITLE OF INVENTION: GENE FOR APRT FROM PLANT TISSUE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,010
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/230,695
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,132
; FILING DATE: 26-MAY-1992
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1811-183 MTS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMEAS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-447-010-9

Query Match 47.5%; Score 28; DB 1; Length 9;
Best Local Similarity 71.4%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 BEYEY 9
Db 1 BEYSLEY 7

RESULT 3
US-08-676-242-3
; Sequence 3, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Michigan
; APPLICANT: Schmaier, Alvin H.
; TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
; FILE REFERENCE: 8820-2 US
; CURRENT APPLICATION NUMBER: US/08/676,242C
; CURRENT FILING DATE: 2000-07-16
; EARLIER APPLICATION NUMBER: 60/000,096
; EARLIER FILING DATE: 1995-06-09
; EARLIER APPLICATION NUMBER: PCT/US96/09940

; EARLIER FILING DATE: 1996-06-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Epitope on
; OTHER INFORMATION: thrombin receptor
US-08-676-242-3

Query Match 47.5%; Score 28; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYE 6
Db 2 PNDKYE 7

RESULT 4
US-09-402-732-3
; Sequence 3, Application US/09402732
; Patent No. 6251855
; GENERAL INFORMATION:
; APPLICANT: Schmaier, Alvin H.
; APPLICANT: Hasan, A.K. Ahmed
; TITLE OF INVENTION: Bradykinin Analogs As Selective Inhibitors of Cell
; TITLE OF INVENTION: Activation
; FILE REFERENCE: 8820-3
; CURRENT APPLICATION NUMBER: US/09/402,732
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 60/046,085
; PRIOR FILING DATE: 1997-04-23
; PRIOR APPLICATION NUMBER: PCT/US98/08015
; PRIOR FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Epitope on
; OTHER INFORMATION: thrombin receptor
US-09-402-732-3

Query Match 47.5%; Score 28; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PNEEYE 6
Db 2 PNDKYE 7

RESULT 5
US-08-318-970B-8
; Sequence 8, Application US/08318970B
; Patent No. 5589573
; GENERAL INFORMATION:
; APPLICANT: Hideaki HAGIWARA, et al.
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
; APPLICANT: Hasan, Ahmed A.K.
; TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
; FILE REFERENCE: 8820-2 US
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Office of Sherman and Shalloway
; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA

ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: Dell System 210; Intel 80 285 Microprocessor
OPERATING SYSTEM: MS DOS 3.3
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,970B
FILING DATE: October 6, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Richard A. Steinberg
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: S-2371
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: H-CDR3-1
OTHER INFORMATION: hypervariable region
US-08-318-970B-8

Query Match 47.5%; Score 28; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EYEYE 8
| | | : |
Db 1 EYDYD 6

RESULT 6
US-08-318-856A-20
; Sequence 20, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,856A
FILING DATE: October 3, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 92 08 068.8
FILING DATE: April 3, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 92 17 704.7
FILING DATE: August 20, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB93/00711
FILING DATE: April 5, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng

REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 263-PP1R1577US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 721-8200
TELEFAX: (202) 721-8250
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-318-856A-20

Query Match 45.8%; Score 27; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNEEYEE 8
| : | |
Db 2 PKDELDYE 9

RESULT 7
US-07-841-997A-6
; Sequence 6, Application US/07841997A
; Patent No. 5422254
; GENERAL INFORMATION:
; APPLICANT: Londeborough, John
; APPLICANT: Vuorio, Outi
; TITLE OF INVENTION: A method to increase the trehalose content
; TITLE OF INVENTION: of organisms by transforming them with the
; TITLE OF INVENTION: structural genes for the short and long chains
; TITLE OF INVENTION: yeast trehalose synthase.
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alko Ltd.
; STREET: PO Box 350
; CITY: Helsinki
; STATE: -
; COUNTRY: Finland
; ZIP: SF-00101
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb
COMPUTER: IBM PC/XT/AT
OPERATING SYSTEM: PC-DOS
SOFTWARE: WP5.1 file exported as DOS text file
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/841,997A
FILING DATE: 19920228
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/836,021
FILING DATE: February 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Mary E. Gormley
REGISTRATION NUMBER: 34409
REFERENCE/DOCKET NUMBER: 920085A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)659-2930
TELEFAX: (202)887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: No
FRAGMENT TYPE: N-terminal
US-07-841-997A-6

Query Match 44.1%; Score 26; DB 1; Length 9;

Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 BEYEV 7
|||:|
Db 3 BEYQY 7

RESULT 8
US-08-290-301-6
; Sequence 6, Application US/08290301
; Patent No. 5792921
; GENERAL INFORMATION:
; APPLICANT: Lonsborough, John
; APPLICANT: Tunnela, Outi
; APPLICANT: Palva, Tupio
; APPLICANT: Holmstrom, Kjell-Ove
; APPLICANT: Wellin, Bjorn
; APPLICANT: Mandel, Abul
; TITLE OF INVENTION: Increasing the trehalose content
; TITLE OF INVENTION: of organisms by transforming them with combinations of
; TITLE OF INVENTION: the structural genes for trehalose synthase.
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alko Ltd.
; STREET: PO Box 350
; CITY: Helsinki
; STATE:
; COUNTRY: Finland
; ZIP: SF-00101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WP5.1 file exported as DOS text file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,301
; FILING DATE: 15 August 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FI 943133
; FILING DATE: 29 June 1994
; APPLICATION NUMBER: PCT/FI93/00049
; FILING DATE: 15 February 1993
; APPLICATION NUMBER: 07/841,997
; FILING DATE: 28 February 1992
; APPLICATION NUMBER: 07/836,021
; FILING DATE: 14 February 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubovcik, Ronald J.
; NAME: Lydon, James C.
; REGISTRATION NUMBER: 25,401
; REFERENCE/DOCKET NUMBER: 30,082
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 467-6300
; TELEFAX: (202) 466-2006
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: Peptide
; HYPOTHETICAL: No
; FRAGMENT TYPE: N-terminal
US-08-290-301-6

Query Match 44.1%; Score 26; DB 1; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 BEYEV 7
|||:|
7

Db 3 BEYQY 7

RESULT 9
US-08-934-222-125
; Sequence 125, Application US/08934222
; Patent No. 5928896
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; TITLE OF INVENTION: Site
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,222
; FILING DATE: 19-SEPT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-934-222-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEVE 6
|||:|
Db 2 PNEENE 7

RESULT 10
US-08-933-402-125
; Sequence 125, Application US/08933402
; Patent No. 5948887
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; TITLE OF INVENTION: Site
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,402
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-933-402-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

RESULT 11
US-09-207-621-125
Sequence 125, Application US/09207621
Patent No. 5952465
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
Restraining Groups Which Flank A Protein-Protein Interaction
NUMBER OF SEQUENCES: 153
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/207,621
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-207-621-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-207-621-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

RESULT 12
US-08-532-818-125
Sequence 125, Application US/08532818
Patent No. 5965698
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
Restraining Groups Which Flank A Protein-Protein Interaction
NUMBER OF SEQUENCES: 153
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-532-818-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

US-08-532-818-125
Sequence 125, Application US/08532818
Patent No. 5965698
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
Restraining Groups Which Flank A Protein-Protein Interaction
NUMBER OF SEQUENCES: 153
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-532-818-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

US-08-532-818-125
Sequence 125, Application US/08532818
Patent No. 5965698
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
Restraining Groups Which Flank A Protein-Protein Interaction
NUMBER OF SEQUENCES: 153
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-532-818-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

US-08-532-818-125
Sequence 125, Application US/08532818
Patent No. 5965698
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
Restraining Groups Which Flank A Protein-Protein Interaction
NUMBER OF SEQUENCES: 153
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-532-818-125

Query Match 44.1%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PNEEYE 6
DB 2 PNEENE 7

Qy 1 PNEEYE 6
Db 2 PNEENE 7

RESULT 13

US-09-231-797-125
; Sequence 125, Application US/09231797
; Patent No. 6084066
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231,797
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-231-797-125

Query Match 44.1%; Score 26; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PNEEYE 6
Db 2 PNEENE 7
RESULT 14
US-08-934-224-125
; Sequence 125, Application US/08934224
; Patent No. 6100044
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,224
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-231-797-125

Query Match 44.1%; Score 26; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYE 6
Db 2 PNEENE 7

RESULT 14

US-08-934-224-125
; Sequence 125, Application US/08934224
; Patent No. 6100044
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,843
; FILING DATE: 19-SEPT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818

ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,224
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-934-224-125

Query Match 44.1%; Score 26; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEYE 6
Db 2 PNEENE 7

RESULT 15

US-08-933-843-125
; Sequence 125, Application US/08933843
; Patent No. 6111069
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,843
; FILING DATE: 19-SEPT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818

;; FILING DATE: 03-MAY-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/143,364
;; FILING DATE: 29-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/051,741
;; FILING DATE: 23-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Isaacson, John P.
;; REGISTRATION NUMBER: 33,751
;; REFERENCE/DOCKET NUMBER: 040433/0148
;; INFORMATION FOR SEQ ID NO: 125:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
US-08-933-843-125

Query Match 44.1%; Score 26; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PNEEVE 6
Db 2 PNEENE 7

Search completed: August 5, 2003, 07:53:02
Job time : 30 secs

THIS PAGE BLANK (USPTO)